An Introduction to Celestial Mechanics

Richard Fitzpatrick

Professor of Physics

The University of Texas at Austin

Contents

1 Introduction **5**

2 Newtonian Mechanics **7**

2.1 Introduction 7
2.2 Newton's Laws of Motion 8
2.3 Newton's First Law of Motion 8
2.4 Newton's Second Law of Motion 11
2.5 Newton's Third Law of Motion 13
2.6 Non-Isolated Systems 16
2.7 Motion in a One-Dimensional Potential 17
2.8 Simple Harmonic Motion 20
2.9 Two-Body Problem 22
2.10 Exercises 23

3 Newtonian Gravity **25**

3.1 Introduction 25
3.2 Gravitational Potential 25
3.3 Gravitational Potential Energy 27
3.4 Axially Symmetric Mass Distributions 28
3.5 Potential Due to a Uniform Sphere 31
3.6 Potential Outside a Uniform Spheroid 32
3.7 Potential Due to a Uniform Ring 35
3.8 Exercises 35

4 Keplerian Orbits **37**

4.1 Introduction 37
4.2 Kepler's Laws 37
4.3 Conservation Laws 38
4.4 Polar Coordinates 38
4.5 Kepler's Second Law 40

4.6 Kepler's First Law . . . 41
4.7 Kepler's Third Law . . . 42
4.8 Orbital Energies . . . 43
4.9 Kepler Problem . . . 45
4.10 Orbital Elements . . . 48
4.11 Binary Star Systems . . . 51
4.12 Exercises . . . 53

5 Orbits in Central Force-Fields 55
5.1 Introduction . . . 55
5.2 Motion in a General Central Force-Field . . . 55
5.3 Motion in a Nearly Circular Orbit . . . 55
5.4 Perihelion Precession of the Planets . . . 58
5.5 Perihelion Precession of Mercury . . . 61
5.6 Exercises . . . 62

6 Rotating Reference Frames 65
6.1 Introduction . . . 65
6.2 Rotating Reference Frames . . . 65
6.3 Centrifugal Acceleration . . . 66
6.4 Coriolis Force . . . 69
6.5 Rotational Flattening of the Earth . . . 71
6.6 Tidal Elongation of the Earth . . . 74
6.7 Tidal Torques . . . 78
6.8 Roche Radius . . . 82
6.9 Exercises . . . 83

7 Lagrangian Mechanics 85
7.1 Introduction . . . 85
7.2 Generalized Coordinates . . . 85
7.3 Generalized Forces . . . 86
7.4 Lagrange's Equation . . . 86
7.5 Generalized Momenta . . . 89
7.6 Exercises . . . 90

8 Rigid Body Rotation 93
8.1 Introduction . . . 93
8.2 Fundamental Equations . . . 93
8.3 Moment of Inertia Tensor . . . 94
8.4 Rotational Kinetic Energy . . . 95
8.5 Principal Axes of Rotation . . . 96
8.6 Euler's Equations . . . 97
8.7 Eulerian Angles . . . 99

8.8 Free Precession of the Earth 102
8.9 McCullough's Formula 103
8.10 Forced Precession and Nutation of the Earth 104
8.11 Exercises 112

9 Three-Body Problem 115
9.1 Introduction 115
9.2 Circular Restricted Three-Body Problem 115
9.3 Jacobi Integral 116
9.4 Tisserand Criterion 117
9.5 Co-Rotating Frame 119
9.6 Lagrange Points 121
9.7 Zero-Velocity Surfaces 124
9.8 Stability of Lagrange Points 127

10 Lunar Motion 133
10.1 Historical Background 133
10.2 Preliminary Analysis 134
10.3 Lunar Equations of Motion 135
10.4 Unperturbed Lunar Motion 138
10.5 Perturbed Lunar Motion 139
10.6 Description of Lunar Motion 144

11 Orbital Perturbation Theory 149
11.1 Introduction 149
11.2 Disturbing Function 149
11.3 Osculating Orbital Elements 150
11.4 Lagrange Brackets 153
11.5 Transformation of Lagrange Brackets 154
11.6 Lagrange's Planetary Equations 158
11.7 Transformation of Lagrange's Equations 160
11.8 Expansion of Lagrange's Equations 161
11.9 Expansion of Disturbing Function 164
11.10 Secular Evolution of Planetary Orbits 172
11.11 Hirayama Families 179

A Vector Algebra and Vector Calculus 185
A.1 Introduction 185
A.2 Scalars and Vectors 185
A.3 Vector Algebra 186
A.4 Cartesian Components of a Vector 187
A.5 Coordinate Transformations 188
A.6 Scalar Product 190

A.7 Vector Product 192
A.8 Rotation 194
A.9 Scalar Triple Product 196
A.10 Vector Triple Product 197
A.11 Vector Calculus 198
A.12 Line Integrals 198
A.13 Vector Line Integrals 201
A.14 Volume Integrals 202
A.15 Gradient 203
A.16 Grad Operator 206
A.17 Curvilinear Coordinates 207
A.18 Exercises 208

B Useful Mathematrics **211**
B.1 Conic Sections 211
B.2 Matrix Eigenvalue Theory 214

1 Introduction

The aim of this book is to bridge the considerable gap between standard undergraduate treatments of celestial mechanics, which rarely advance much beyond two-body orbit theory, and full-blown graduate treatments, such as that by Murray & Dermott. The material presented here is intended to be intelligible to advanced undergraduate and beginning graduate students. A knowledge of elementary Newtonian mechanics is assumed. However, those non-elementary topics in mechanics that are needed to account for the motions of celestial bodies (*e.g.*, gravitational potential theory, motion in rotating reference frames, Lagrangian mechanics, Eulerian rigid body rotation theory) are derived in the text. This is entirely appropriate, since these many of these topics were originally developed in order to investigate specific problems in celestial mechanics (often, the very problems that they are used to examine in this book.) It is taken for granted that readers are familiar with the fundamentals of integral and differential calculus, ordinary differential equations, and linear algebra. On the other hand, vector analysis plays such a central role in the study of celestial motion that a brief, but fairly comprehensive, review of this subject area is provided in Appendix A.

Celestial mechanics is the branch of astronomy that is concerned with the motions of celestial objects—in particular, the objects that make up the Solar System. The main aim of celestial mechanics is to reconcile these motions with the predictions of Newtonian mechanics. Modern analytic celestial mechanics started in 1687 with the publication of the Principia by Isaac Newton (1643–1727), and was subsequently developed into a mature science by celebrated scientists such as Euler (1707–1783), Clairaut (1713–1765), D'Alembert (1717–1783), Lagrange (1736–1813), Laplace (1749–1827), and Gauss (1777–1855). This book is largely devoted to the study of the "classical" problems of celestial mechanics that were investigated by these scientists. These problems include the figure of the Earth, the tidal interaction between the Earth and the Moon, the free and forced precession and nutation of the Earth, the three-body problem, the orbit of the Moon, and the effect of interplanetary gravitational interactions on planetary orbits. This book does not discuss positional astronomy: *i.e.*, the branch of astronomy that is concerned with finding the positions of celestial objects in the Earth's sky at a particular instance in time. Nor does this book discuss the relatively modern (but extremely complicated) developments in celestial mechanics engendered by the advent of cheap fast computers, as well as data from unmanned space probes: *e.g.*, the resonant interaction of planetary moons, chaotic motion, the dynamics of planetary rings. Interested readers are directed to the book by Murray & Dermott.

The major sources for the material appearing in this book are as follows:

An Elementary Treatise on the Lunar Theory, H. Godfray (Macmillan, London UK, 1853).

An Introductory Treatise on the Lunar Theory, E.W. Brown (Cambridge University Press, Cambridge UK, 1896).

Lectures on the Lunar Theory, J.C. Adams (Cambridge University Press, Cambridge UK, 1900).

An Introduction to Celestial Mechanics, F.R. Moulton, 2nd Revised Edition (Macmillan, New York NY, 1914).

Dynamics, H. Lamb, 2nd Edition (Cambridge University Press, Cambridge UK, 1923).

Celestial Mechanics, W.M. Smart (Longmans, London UK, 1953).

An Introductory Treatise on Dynamical Astronomy, H.C. Plummer (Dover, New York NY, 1960).

Methods of Celestial Mechanics, D. Brouwer, and G.M. Clemence (Academic Press, New York NY, 1961).

Celestial Mechanics: Volume II, Part 1: Perturbation Theory, Y. Hagihara (MIT Press, Cambridge MA, 1971).

Analytical Mechanics, G.R. Fowles (Holt, Rinehart, and Winston, New York NY, 1977).

Orbital Motion, A.E. Roy (Wiley, New York NY, 1978).

Solar System Dynamics, C.D. Murray, and S.F. Dermott (Cambridge University Press, Cambridge UK, 1999).

Classical Mechanics, 3rd Edition, H. Goldstein, C. Poole, and J. Safko (Addison-Wesley, San Fransisco CA, 2002).

Classical Dynamics of Particles and Systems, 5th Edition, S.T. Thornton, and J.B. Marion (Brooks/Cole—Thomson Learning, Belmont CA, 2004).

Analytical Mechanics, 7th Edition, G.R. Fowles, and G.L. Cassiday (Brooks/Cole—Thomson Learning, Belmont CA, 2005).

Astronomical Algorithms, J. Meeus, 2nd Edition (Willmann-Bell, Richmond VA, 2005).

2 Newtonian Mechanics

2.1 Introduction

Newtonian mechanics is a mathematical model whose purpose is to account for the motions of the various objects in the Universe. The general principles of this model were first enunciated by Sir Isaac Newton in a work entitled *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy). This work, which was published in 1687, is nowadays more commonly referred to as the *Principa*.[1]

Up until the beginning of the 20th century, Newtonian mechanics was thought to constitute a *complete* description of all types of motion occurring in the Universe. We now know that this is not the case. The modern view is that Newton's model is only an *approximation* which is valid under certain circumstances. The model breaks down when the velocities of the objects under investigation approach the speed of light in vacuum, and must be modified in accordance with Einstein's *special theory of relativity*. The model also fails in regions of space which are sufficiently curved that the propositions of Euclidean geometry do not hold to a good approximation, and must be augmented by Einstein's *general theory of relativity*. Finally, the model breaks down on atomic and subatomic length-scales, and must be replaced by *quantum mechanics*. In this book, we shall neglect relativistic and quantum effects altogether. It follows that we must restrict our investigations to the motions of *large* (compared to an atom) *slow* (compared to the speed of light) objects moving in *Euclidean* space. Fortunately, virtually all of the motions encountered in celestial mechanics fall into this category.

Newton very deliberately modeled his approach in the Principia on that taken in *Euclid's Elements*. Indeed, Newton's theory of motion has much in common with a conventional *axiomatic system* such as Euclidean geometry. Like all axiomatic systems, Newtonian mechanics starts from a set of terms that are *undefined* within the system. In this case, the fundamental terms are *mass*, *position*, *time*, and *force*. It is taken for granted that we understand what these terms mean, and, furthermore, that they correspond to *measurable* quantities which can be ascribed to, or associated with, objects in the world around us. In particular, it is assumed that the ideas of position in space, distance in space, and position as a function of time in space, are correctly described by the Euclidean vector algebra and vector calculus discussed in Appendix A. The next component of an axiomatic system is a set of *axioms*. These are a set of *unproven* propositions, involving the undefined terms, from which all other propositions in the system can be derived via logic and mathematical analysis. In the present case, the axioms are called *Newton's laws of motion*, and can only be justified via experimental observation. Note, incidentally, that Newton's laws, in their

[1]An excellent discussion of the historical development of Newtonian mechanics, as well as the physical and philosophical assumptions which underpin this theory, is given in *The Discovery of Dynamics: A Study from a Machian Point of View of the Discovery and the Structure of Dynamical Theories*, J.B. Barbour (Oxford University Press, Oxford UK, 2001).

primitive form, are only applicable to *point objects* (*i.e.*, objects of negligible spatial extent). However, these laws can be applied to extended objects by treating them as collections of point objects.

One difference between an axiomatic system and a physical theory is that, in the latter case, even if a given prediction has been shown to follow necessarily from the axioms of the theory, it is still incumbent upon us to test the prediction against experimental observations. Lack of agreement might indicate faulty experimental data, faulty application of the theory (for instance, in the case of Newtonian mechanics, there might be forces at work which we have not identified), or, as a last resort, incorrectness of the theory. Fortunately, Newtonian mechanics has been found to give predictions that are in excellent agreement with experimental observations in all situations in which it would be expected to hold.

In the following, it is assumed that we know how to set up a rigid *Cartesian frame of reference*, and how to measure the positions of point objects as functions of time within that frame. It is also taken for granted that we have some basic familiarity with the laws of mechanics, and with standard mathematics up to, and including, calculus, as well as the vector analysis outlined in Appendix A.

2.2 Newton's Laws of Motion

Newton's laws of motion, in the rather obscure language of the Principia, take the following form:

1. Every body continues in its state of rest, or uniform motion in a straight-line, unless compelled to change that state by forces impressed upon it.

2. The change of motion (*i.e.*, momentum) of an object is proportional to the force impressed upon it, and is made in the direction of the straight-line in which the force is impressed.

3. To every action there is always opposed an equal reaction; or, the mutual actions of two bodies upon each other are always equal and directed to contrary parts.

Let us now examine how these laws can be applied to a system of point objects.

2.3 Newton's First Law of Motion

Newton's first law of motion essentially states that a point object subject to zero net external force moves in a straight-line with a constant speed (*i.e.*, it does not accelerate). However, this is only true in special frames of reference called *inertial frames*. Indeed, we can think of Newton's first law as the *definition* of an inertial frame: *i.e.*, an inertial frame of reference is one in which a point object subject to zero net external force moves in a straight-line with constant speed.

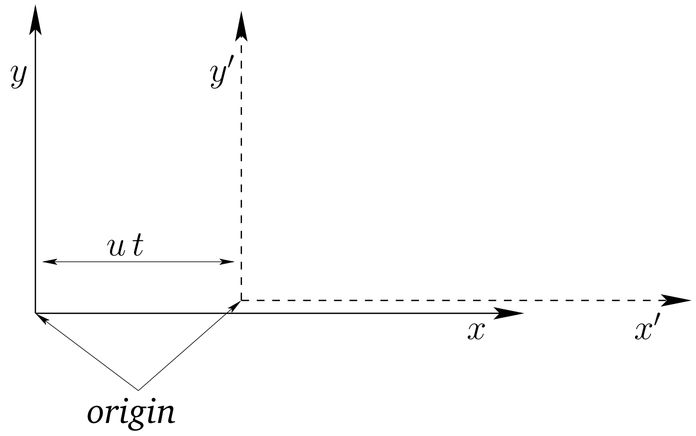


Figure 2.1: *A Galilean coordinate transformation.*

Suppose that we have found an inertial frame of reference. Let us set up a Cartesian coordinate system in this frame. The motion of a point object can now be specified by giving its position vector, $\mathbf{r} \equiv (x,\, y,\, z)$, with respect to the origin of the coordinate system, as a function of time, t. Consider a second frame of reference moving with some *constant* velocity $\mathbf{u}$ with respect to the first frame. Without loss of generality, we can suppose that the Cartesian axes in the second frame are parallel to the corresponding axes in the first frame, that $\mathbf{u} \equiv (u, 0, 0)$, and, finally, that the origins of the two frames instantaneously coincide at $t = 0$—see Figure 2.1. Suppose that the position vector of our point object is $\mathbf{r}' \equiv (x',\, y',\, z')$ in the second frame of reference. It is evident, from Figure 2.1, that at any given time, t, the coordinates of the object in the two reference frames satisfy

$$x' = x - u\,t, \tag{2.1}$$

$$y' = y, \tag{2.2}$$

$$z' = z. \tag{2.3}$$

This coordinate transformation was first discovered by Galileo Galilei, and is known as a *Galilean transformation* in his honor.

By definition, the instantaneous velocity of the object in our first reference frame is given by $\mathbf{v} = d\mathbf{r}/dt \equiv (dx/dt,\, dy/dt,\, dz/dt)$, with an analogous expression for the velocity, $\mathbf{v}'$, in the second frame. It follows from differentiation of Equations (2.1)–(2.3) with respect to time that the velocity components in the two frames satisfy

$$v_x' = v_x - u, \tag{2.4}$$

$$v_y' = v_y, \tag{2.5}$$

$$v_z' = v_z. \tag{2.6}$$

These equations can be written more succinctly as

$$\mathbf{v}' = \mathbf{v} - \mathbf{u}. \tag{2.7}$$

Finally, by definition, the instantaneous acceleration of the object in our first reference frame is given by $\mathbf{a} = d\mathbf{v}/dt \equiv (dv_x/dt,\, dv_y/dt,\, dv_z/dt)$, with an analogous expression for the acceleration, $\mathbf{a}'$, in the second frame. It follows from differentiation of Equations (2.4)–(2.6) with respect to time that the acceleration components in the two frames satisfy

$$a_x' = a_x, \tag{2.8}$$

$$a_y' = a_y, \tag{2.9}$$

$$a_z' = a_z. \tag{2.10}$$

These equations can be written more succinctly as

$$\mathbf{a}' = \mathbf{a}. \tag{2.11}$$

According to Equations (2.7) and (2.11), if an object is moving in a straight-line with constant speed in our original inertial frame (*i.e.*, if $\mathbf{a} = \mathbf{0}$) then it also moves in a (different) straight-line with (a different) constant speed in the second frame of reference (*i.e.*, $\mathbf{a}' = \mathbf{0}$). Hence, we conclude that the second frame of reference is *also* an inertial frame.

A simple extension of the above argument allows us to conclude that there are an *infinite* number of different inertial frames moving with *constant velocities* with respect to one another. Newton throught that one of these inertial frames was special, and defined an absolute standard of rest: *i.e.*, a static object in this frame was in a state of absolute rest. However, Einstein showed that this is not the case. In fact, there is no absolute standard of rest: *i.e.*, all motion is relative—hence, the name "relativity" for Einstein's theory. Consequently, one inertial frame is just as good as another as far as Newtonian mechanics is concerned.

But, what happens if the second frame of reference *accelerates* with respect to the first? In this case, it is not hard to see that Equation (2.11) generalizes to

$$\mathbf{a}' = \mathbf{a} - \frac{d\mathbf{u}}{dt}, \tag{2.12}$$

where $\mathbf{u}(t)$ is the instantaneous velocity of the second frame with respect to the first. According to the above formula, if an object is moving in a straight-line with constant speed in the first frame (*i.e.*, if $\mathbf{a} = \mathbf{0}$) then it does not move in a straight-line with constant speed in the second frame (*i.e.*, $\mathbf{a}' \neq \mathbf{0}$). Hence, if the first frame is an inertial frame then the second is *not*.

A simple extension of the above argument allows us to conclude that any frame of reference which accelerates with respect to a given inertial frame is not itself an inertial frame.

For most practical purposes, when studying the motions of objects close to the Earth's surface, a reference frame that is fixed with respect to this surface is approximately inertial. However, if the trajectory of a projectile within such a frame is measured to high precision then it will be found to deviate slightly from the predictions of Newtonian mechanics—see Chapter 6. This deviation is due to the fact that the Earth is rotating, and its surface is

therefore accelerating towards its axis of rotation. When studying the motions of objects in orbit around the Earth, a reference frame whose origin is the center of the Earth, and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. However, if such orbits are measured to extremely high precision then they will again be found to deviate very slightly from the predictions of Newtonian mechanics. In this case, the deviation is due to the Earth's orbital motion about the Sun. When studying the orbits of the planets in the Solar System, a reference frame whose origin is the center of the Sun, and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. In this case, any deviations of the orbits from the predictions of Newtonian mechanics due to the orbital motion of the Sun about the galactic center are far too small to be measurable. It should be noted that it is impossible to identify an *absolute* inertial frame—the best approximation to such a frame would be one in which the cosmic microwave background appears to be (approximately) isotropic. However, for a given dynamical problem, it is always possible to identify an *approximate* inertial frame. Furthermore, any deviations of such a frame from a true inertial frame can be incorporated into the framework of Newtonian mechanics via the introduction of so-called fictitious forces—see Chapter 6.

2.4 Newton's Second Law of Motion

Newton's second law of motion essentially states that if a point object is subject to an external force, $\mathbf{f}$, then its equation of motion is given by

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}, \tag{2.13}$$

where the momentum, $\mathbf{p}$, is the product of the object's inertial mass, m, and its velocity, $\mathbf{v}$. If m is not a function of time then the above expression reduces to the familiar equation

$$m\,\frac{d\mathbf{v}}{dt} = \mathbf{f}. \tag{2.14}$$

Note that this equation is only valid in a *inertial frame*. Clearly, the inertial mass of an object measures its reluctance to deviate from its preferred state of uniform motion in a straight-line (in an inertial frame). Of course, the above equation of motion can only be solved if we have an independent expression for the force, $\mathbf{f}$ (*i.e.*, a law of force). Let us suppose that this is the case.

An important corollary of Newton's second law is that force is a *vector quantity*. This must be the case, since the law equates force to the product of a scalar (mass) and a vector (acceleration). Note that acceleration is obviously a vector because it is directly related to displacement, which is the prototype of all vectors—see Appendix A. One consequence of force being a vector is that two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, both acting at a given point, have the same effect as a single force, $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, acting at the same point, where the summation is performed according to the laws of vector addition—see Appendix A. Likewise, a single force, $\mathbf{f}$, acting at a given point, has the same effect as two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, acting at the

same point, provided that $\mathbf{f}_1 + \mathbf{f}_2 = \mathbf{f}$. This method of combining and splitting forces is known as the *resolution of forces*, and lies at the heart of many calculations in Newtonian mechanics.

Taking the scalar product of Equation (2.14) with the velocity, $\mathbf{v}$, we obtain

$$m\,\mathbf{v}\cdot\frac{d\mathbf{v}}{dt} = \frac{m}{2}\frac{d(\mathbf{v}\cdot\mathbf{v})}{dt} = \frac{m}{2}\frac{dv^2}{dt} = \mathbf{f}\cdot\mathbf{v}. \tag{2.15}$$

This can be written

$$\frac{dK}{dt} = \mathbf{f}\cdot\mathbf{v}. \tag{2.16}$$

where

$$K = \frac{1}{2}\,m\,v^2. \tag{2.17}$$

The right-hand side of Equation (2.16) represents the rate at which the force does work on the object (see Section A.6): *i.e.*, the rate at which the force transfers energy to the object. The quantity K represents the energy that the object possesses by virtue of its motion. This type of energy is generally known as *kinetic energy*. Thus, Equation (2.16) states that any work done on point object by an external force goes to increase the object's kinetic energy.

Suppose that, under the action of the force, $\mathbf{f}$, our object moves from point P at time t_1 to point Q at time t_2. The net change in the object's kinetic energy is obtained by integrating Equation (2.16):

$$\Delta K = \int_{t_1}^{t_2} \mathbf{f}\cdot\mathbf{v}\,dt = \int_P^Q \mathbf{f}\cdot d\mathbf{r}, \tag{2.18}$$

since $\mathbf{v} = d\mathbf{r}/dt$. Here, $d\mathbf{r}$ is an element of the object's path between points P and Q, and the integral in $\mathbf{r}$ represents the net work done by the force as the object moves along the path from P to Q.

As described in Section A.15, there are basically two kinds of forces in nature. Firstly, those for which line integrals of the type $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ depend on the end points, but not on the path taken between these points. Secondly, those for which line integrals of the type $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ depend both on the end points, and the path taken between these points. The first kind of force is termed *conservative*, whereas the second kind is termed *non-conservative*. It is also demonstrated in Section A.15 that if the line integral $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ is *path independent* then the force $\mathbf{f}$ can always be written as the gradient of a scalar field. In other words, all conservative forces satisfy

$$\mathbf{f}(\mathbf{r}) = -\nabla U, \tag{2.19}$$

for some scalar field $U(\mathbf{r})$. Note that

$$\int_P^Q \nabla U\cdot d\mathbf{r} \equiv \Delta U = U(Q) - U(P), \tag{2.20}$$

irrespective of the path taken between P and Q. Hence, it follows from Equation (2.18) that

$$\Delta K = -\Delta U \tag{2.21}$$

for conservative forces. Another way of writing this is

$$E = K + U = \text{constant}. \tag{2.22}$$

Of course, we recognize this as an *energy conservation equation*: E is the object's total energy, which is conserved; K is the energy the object has by virtue of its motion, otherwise know as its *kinetic energy*; and U is the energy the object has by virtue of its position, otherwise known as its *potential energy*. Note, however, that we can only write such energy conservation equations for conservative forces. Gravity is an obvious example of a conservative force. Non-conservative forces, on the other hand, do not conserve energy. In general, this is because of some sort of frictional energy loss. Note that potential energy is undefined to an arbitrary additive constant. In fact, it is only the *difference* in potential energy between different points in space that is well-defined.

2.5 Newton's Third Law of Motion

Consider a system of N mutually interacting point objects. Let the ith object, whose mass is m_i, be located at vector displacement $\mathbf{r}_i$. Suppose that this object exerts a force $\mathbf{f}_{ji}$ on the jth object. Likewise, suppose that the jth object exerts a force $\mathbf{f}_{ij}$ on the ith object. Newton's third law of motion essentially states that these two forces are equal and opposite, irrespective of their nature. In other words,

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}. \tag{2.23}$$

One corollary of Newton's third law is that an object cannot exert a force on itself. Another corollary is that all forces in the Universe have corresponding reactions. The only exceptions to this rule are the fictitious forces which arise in non-inertial reference frames (*e.g.*, the centrifugal and Coriolis forces which appear in rotating reference frames—see Chapter 6). Fictitious forces do not possess reactions.

It should be noted that Newton's third law implies *action at a distance*. In other words, if the force that object i exerts on object j suddenly changes then Newton's third law demands that there must be an *immediate* change in the force that object j exerts on object i. Moreover, this must be true irrespective of the distance between the two objects. However, we now know that Einstein's theory of relativity forbids information from traveling through the Universe faster than the velocity of light in vacuum. Hence, action at a distance is also forbidden. In other words, if the force that object i exerts on object j suddenly changes then there must be a *time delay*, which is at least as long as it takes a light ray to propagate between the two objects, before the force that object j exerts on object i can respond. Of course, this means that Newton's third law is not, strictly speaking, correct. However, as long as we restrict our investigations to the motions of dynamical systems on time-scales that are long compared to the time required for light-rays to traverse these systems, Newton's third law can be regarded as being approximately correct.

In an inertial frame, Newton's second law of motion applied to the ith object yields

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{j=1,N}^{j\neq i} \mathbf{f}_{ij}. \tag{2.24}$$

Note that the summation on the right-hand side of the above equation excludes the case $j = i$, since the ith object cannot exert a force on itself. Let us now take the above equation and sum it over all objects. We obtain

$$\sum_{i=1,N} m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{i,j=1,N}^{j\neq i} \mathbf{f}_{ij}. \tag{2.25}$$

Consider the sum over forces on the right-hand side of the above equation. Each element of this sum—$\mathbf{f}_{ij}$, say—can be paired with another element—$\mathbf{f}_{ji}$, in this case—which is equal and opposite, according to Newton's third law. In other words, the elements of the sum all cancel out in pairs. Thus, the net value of the sum is *zero*. It follows that the above equation can be written

$$M \frac{d^2\mathbf{r}_{cm}}{dt^2} = \mathbf{0}, \tag{2.26}$$

where $M = \sum_{i=1}^{N} m_i$ is the total mass. The quantity $\mathbf{r}_{cm}$ is the vector displacement of the *center of mass* of the system, which is an imaginary point whose coordinates are the mass weighted averages of the coordinates of the objects that constitute the system: *i.e.*,

$$\mathbf{r}_{cm} = \frac{\sum_{i=1}^{N} m_i\,\mathbf{r}_i}{\sum_{i=1}^{N} m_i}. \tag{2.27}$$

According to Equation (2.26), the center of mass of the system moves in a uniform straight-line, in accordance with Newton's first law of motion, irrespective of the nature of the forces acting between the various components of the system.

Now, if the center of mass moves in a uniform straight-line then the center of mass velocity,

$$\frac{d\mathbf{r}_{cm}}{dt} = \frac{\sum_{i=1}^{N} m_i\,d\mathbf{r}_i/dt}{\sum_{i=1}^{N} m_i}, \tag{2.28}$$

is a constant of the motion. However, the momentum of the ith object takes the form $\mathbf{p}_i = m_i\,d\mathbf{r}_i/dt$. Hence, the total momentum of the system is written

$$\mathbf{P} = \sum_{i=1}^{N} m_i \frac{d\mathbf{r}_i}{dt}. \tag{2.29}$$

A comparison of Equations (2.28) and (2.29) suggests that $\mathbf{P}$ is also a constant of the motion. In other words, the total momentum of the system is a *conserved* quantity, irrespective

of the nature of the forces acting between the various components of the system. This result (which only holds if there is no net external force acting on the system) is a direct consequence of Newton's third law of motion.

Taking the vector product of Equation (2.24) with the position vector $\mathbf{r}_i$, we obtain

$$m_i\,\mathbf{r}_i\times\frac{d^2\mathbf{r}_i}{dt^2}=\sum_{j=1,N}^{j\neq i}\mathbf{r}_i\times\mathbf{f}_{ij}. \tag{2.30}$$

However, it is easily seen that

$$m_i\,\mathbf{r}_i\times\frac{d^2\mathbf{r}_i}{dt^2}=\frac{d}{dt}\left(m_i\,\mathbf{r}_i\times\frac{d\mathbf{r}_i}{dt}\right)=\frac{d\mathbf{l}_i}{dt}, \tag{2.31}$$

where

$$\mathbf{l}_i=m_i\,\mathbf{r}_i\times\frac{d\mathbf{r}_i}{dt} \tag{2.32}$$

is the *angular momentum* of the ith particle about the origin of our coordinate system. The total angular momentum of the system (about the origin) takes the form

$$\mathbf{L}=\sum_{i=1,N}\mathbf{l}_i \tag{2.33}$$

Hence, summing Equation (2.30) over all particles, we obtain

$$\frac{d\mathbf{L}}{dt}=\sum_{i,j=1,N}^{i\neq j}\mathbf{r}_i\times\mathbf{f}_{ij}. \tag{2.34}$$

Consider the sum on the right-hand side of the above equation. A general term, $\mathbf{r}_i\times\mathbf{f}_{ij}$, in this sum can always be paired with a matching term, $\mathbf{r}_j\times\mathbf{f}_{ji}$, in which the indices have been swapped. Making use of Equation (2.23), the sum of a general matched pair can be written

$$\mathbf{r}_i\times\mathbf{f}_{ij}+\mathbf{r}_j\times\mathbf{f}_{ji}=(\mathbf{r}_i-\mathbf{r}_j)\times\mathbf{f}_{ij}. \tag{2.35}$$

Let us assume that the forces acting between the various components of the system are *central* in nature, so that $\mathbf{f}_{ij}$ is parallel to $\mathbf{r}_i-\mathbf{r}_j$. In other words, the force exerted on object j by object i either points directly toward, or directly away from, object i, and *vice versa*. This is a reasonable assumption, since virtually all of the forces that we encounter in celestial mechanics are of this type (*e.g.*, gravity). It follows that if the forces are central then the vector product on the right-hand side of the above expression is zero. We conclude that

$$\mathbf{r}_i\times\mathbf{f}_{ij}+\mathbf{r}_j\times\mathbf{f}_{ji}=\mathbf{0}, \tag{2.36}$$

for all values of i and j. Thus, the sum on the right-hand side of Equation (2.34) is zero for any kind of central force. We are left with

$$\frac{d\mathbf{L}}{dt}=\mathbf{0}. \tag{2.37}$$

In other words, the total angular momentum of the system is a *conserved* quantity, provided that the different components of the system interact via *central* forces (and there is no net external torque acting on the system).

2.6 Non-Isolated Systems

Up to now, we have only considered *isolated* dynamical systems, in which all of the forces acting on the system originate from within the system itself. Let us now generalize our approach to deal with *non-isolated* dynamical systems, in which some of the forces originate outside the system. Consider a system of N mutually interacting point objects. Let m_i and $\mathbf{r}_i$ be the mass and position vector of the ith object, respectively. Suppose that the ith object is subject to two forces. First, an *internal force* which originates from the other objects in the system, and second an *external force* which originates outside the system. In other words, let the force acting on the ith object take the form

$$\mathbf{f}_i = \sum_{j=1,N}^{j\neq i} \mathbf{f}_{ij} + \mathbf{F}_i, \tag{2.38}$$

where $\mathbf{f}_{ij}$ is the internal force exerted by object j on object i, and $\mathbf{F}_i$ the net external force acting on object i.

The equation of motion of the ith object is

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{f}_i = \sum_{j=1,N}^{j\neq i} \mathbf{f}_{ij} + \mathbf{F}_i. \tag{2.39}$$

Summing over all objects, we obtain

$$\sum_{i=1,N} m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{i,j=1,N}^{j\neq i} \mathbf{f}_{ij} + \sum_{i=1,N} \mathbf{F}_i, \tag{2.40}$$

which reduces to

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}, \tag{2.41}$$

where

$$\mathbf{F} = \sum_{i=1,N} \mathbf{F}_i \tag{2.42}$$

is the net external force acting on the system. Here, the sum over the internal forces has cancelled out in pairs due to Newton's third law of motion (see Section 2.5). We conclude that if there is a net external force acting on the system then the total linear momentum evolves in time according to the simple equation (2.41), but is completely unaffected by any internal forces. The fact that Equation (2.41) is similar in form to Equation (2.13) suggests that the center of mass of a system consisting of many point objects has analogous

dynamics to a single point object, whose mass is the total system mass, moving under the action of the net external force.

Taking $\mathbf{r}_i\times$ Equation (2.39), and summing over all objects, we obtain

$$\frac{d\mathbf{L}}{dt} = \mathbf{T}, \tag{2.43}$$

where

$$\mathbf{T} = \sum_{i=1,N} \mathbf{r}_i \times \mathbf{F}_i \tag{2.44}$$

is the net external torque acting on the system. Here, the sum over the internal torques has cancelled out in pairs, assuming that the internal forces are central in nature (see Section 2.5). We conclude that if there is a net external torque acting on the system then the total angular momentum evolves in time according to the simple equation (2.43), but is completely unaffected by any internal torques.

2.7 Motion in a One-Dimensional Potential

As a simple illustration of the application of Newton's laws of motion, consider a point particle of mass m moving in the x-direction, say, under the action of some x-directed force $f(x)$. Suppose that $f(x)$ is a conservative force: *e.g.*, gravity. In this case, according to Equation (2.19), we can write

$$f(x) = -\frac{dU(x)}{dx}, \tag{2.45}$$

where $U(x)$ is the potential energy of the particle at position x.

Let the curve $U(x)$ take the form shown in Figure 2.2. For instance, this curve might represent the gravitational potential energy of a cyclist freewheeling in a hilly region. Observe that we have set the potential energy at infinity to zero (which we are generally free to do, since potential energy is undefined to an arbitrary additive constant). This is a fairly common convention. What can we deduce about the motion of the particle in this potential?

Well, we know that the total energy, E—which is the sum of the kinetic energy, K, and the potential energy, U—is a *constant* of the motion—see Equation (2.22). Hence, we can write

$$K(x) = E - U(x). \tag{2.46}$$

However, we also know that a kinetic energy can never be negative [since $K = (1/2)\, m\, v^2$, and neither m nor v^2 can be negative]. Hence, the above expression tells us that the particle's motion is restricted to the region (or regions) in which the potential energy curve $U(x)$ falls below the value E. This idea is illustrated in Figure 2.2. Suppose that the total energy of the system is E_0. It is clear, from the figure, that the particle is trapped inside one or other of the two dips in the potential—these dips are generally referred to as *potential*

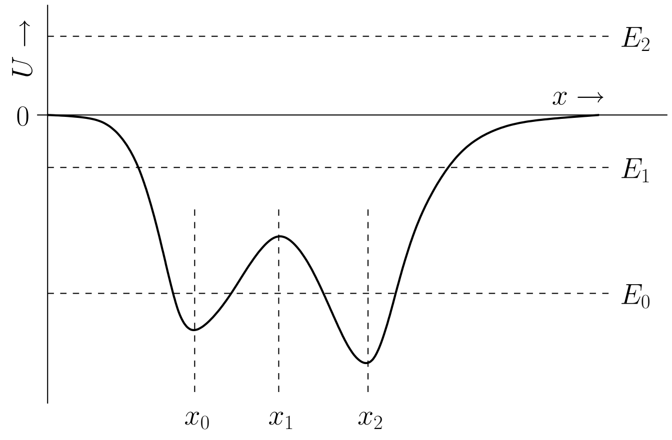


Figure 2.2: *A potential energy curve.*

wells. Suppose that we now raise the energy to E_1. In this case, the particle is free to enter or leave each of the potential wells, but its motion is still *bounded* to some extent, since it clearly cannot move off to infinity. Finally, let us raise the energy to E_2. Now the particle is *unbounded*: *i.e.*, it can move off to infinity. In conservative systems in which it makes sense to adopt the convention that the potential energy at infinity is zero, bounded systems are characterized by $E < 0$, whereas unbounded systems are characterized by $E > 0$.

The above discussion suggests that the motion of a particle moving in a potential generally becomes less bounded as the total energy E of the system increases. Conversely, we would expect the motion to become more bounded as E decreases. In fact, if the energy becomes sufficiently small then it appears likely that the system will settle down in some *equilibrium state* in which the particle remains stationary. Let us try to identify any prospective equilibrium states in Figure 2.2. If the particle remains stationary then it must be subject to zero force (otherwise it would accelerate). Hence, according to Equation (2.45), an equilibrium state is characterized by

$$\frac{dU}{dx} = 0. \tag{2.47}$$

In other words, a equilibrium state corresponds to either a *maximum* or a *minimum* of the potential energy curve $U(x)$. It can be seen that the $U(x)$ curve shown in Figure 2.2 has three associated equilibrium states located at $x = x_0$, $x = x_1$, and $x = x_2$.

Let us now make a distinction between *stable* equilibrium points and *unstable* equilibrium points. When the particle is slightly displaced from a stable equilibrium point then the resultant force acting on it must always be such as to return it to this point. In other words, if $x = x_0$ is an equilibrium point then we require

$$\left.\frac{df}{dx}\right|_{x_0} < 0 \tag{2.48}$$

for stability: *i.e.*, if the particle is displaced to the right, so that $x - x_0 > 0$, then the force

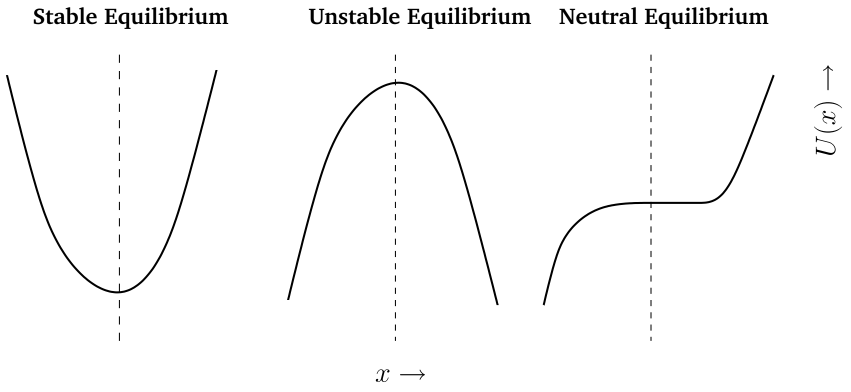


Figure 2.3: *Different types of equilibrium point.*

must act to the left, so that $f < 0$, and *vice versa*. Likewise, if

$$\left.\frac{df}{dx}\right|_{x_0} > 0 \tag{2.49}$$

then the equilibrium point $x = x_0$ is unstable. It follows, from Equation (2.45), that stable equilibrium points are characterized by

$$\frac{d^2U}{dx^2} > 0. \tag{2.50}$$

In other words, a stable equilibrium point corresponds to a *minimum* of the potential energy curve $U(x)$. Likewise, an unstable equilibrium point corresponds to a *maximum* of the $U(x)$ curve. Hence, we conclude that, in Figure 2.2, $x = x_0$ and $x = x_2$ are stable equilibrium points, whereas $x = x_1$ is an unstable equilibrium point. Of course, this makes perfect sense if we think of $U(x)$ as a gravitational potential energy curve, so that U is directly proportional to height. In this case, all we are saying is that it is easy to confine a low energy mass at the bottom of a smooth valley, but very difficult to balance the same mass on the top of a smooth hill (since any slight displacement of the mass will cause it to slide down the hill). Note, finally, that if

$$\frac{dU}{dx} = \frac{d^2U}{dx^2} = 0 \tag{2.51}$$

at any point (or in any region) then we have what is known as a *neutral equilibrium* point. We can move the particle slightly away from such a point and it will still remain in equilibrium (*i.e.*, it will neither attempt to return to its initial state, nor will it continue to move). A neutral equilibrium point corresponds to a *flat spot* in a $U(x)$ curve. See Figure 2.3.

The equation of motion of a particle moving in one dimension under the action of a conservative force is, in principle, integrable. Since $K = (1/2)\, m v^2$, the energy conserva-

tion equation (2.46) can be rearranged to give

$$v = \pm\left(\frac{2\,[E-U(x)]}{m}\right)^{1/2}, \tag{2.52}$$

where the $\pm$ signs correspond to motion to the left and to the right, respectively. However, since $v = dx/dt$, this expression can be integrated to give

$$t = \pm\left(\frac{m}{2\,E}\right)^{1/2}\int_{x_0}^{x}\frac{dx'}{\sqrt{1-U(x')/E}}, \tag{2.53}$$

where $x(t=0) = x_0$. For sufficiently simple potential functions, $U(x)$, the above equation can be solved to give x as a function of t. For instance, if $U = (1/2)\,k\,x^2$, $x_0 = 0$, and the plus sign is chosen, then

$$t = \left(\frac{m}{k}\right)^{1/2}\int_0^{(k/2\,E)^{1/2}\,x}\frac{dy}{\sqrt{1-y^2}} = \left(\frac{m}{k}\right)^{1/2}\sin^{-1}\left(\left[\frac{k}{2\,E}\right]^{1/2} x\right), \tag{2.54}$$

which can be inverted to give

$$x = a\,\sin(\omega\,t), \tag{2.55}$$

where $a = \sqrt{2\,E/k}$ and $\omega = \sqrt{k/m}$. Note that the particle reverses direction each time it reaches one of the so-called *turning points* ($x = \pm a$) at which $U = E$ (and, so $K = 0$).

2.8 Simple Harmonic Motion

Consider the motion of a point particle of mass m, moving in one dimension, that is slightly displaced from a *stable* equilibrium point located at $x = 0$. Suppose that the particle is moving in the conservative force-field $f(x)$. According to the above analysis, in order for $x = 0$ to correspond to a stable equilibrium point we require both

$$f(0) = 0, \tag{2.56}$$

and

$$\frac{df(0)}{dx} < 0. \tag{2.57}$$

Now, our particle obeys Newton's second law of motion,

$$m\,\frac{d^2x}{dt^2} = f(x). \tag{2.58}$$

Let us assume that the particle always stays fairly close to its equilibrium point. In this case, to a good approximation, we can represent $f(x)$ via a truncated Taylor expansion about this point. In other words,

$$f(x) \simeq f(0) + \frac{df(0)}{dx}\,x + \mathcal{O}(x^2). \tag{2.59}$$

However, according to (2.56) and (2.57), the above expression can be written

$$f(x) \simeq -m\,\omega_0^2\,x, \tag{2.60}$$

where $df(0)/dx = -m\,\omega_0^2$. Hence, we conclude that our particle satisfies the following approximate equation of motion,

$$\frac{d^2x}{dt^2} + \omega_0^2\,x \simeq 0, \tag{2.61}$$

provided that it does not stray too far from its equilibrium point: *i.e.*, provided $|x|$ does not become too large.

Equation (2.61) is called the *simple harmonic equation*, and governs the motion of *all* one-dimensional conservative systems which are slightly perturbed from some stable equilibrium state. The solution of Equation (2.61) is well-known:

$$x(t) = a\,\sin(\omega_0\,t - \phi_0). \tag{2.62}$$

The pattern of motion described by above expression, which is called *simple harmonic motion*, is *periodic* in time, with repetition period $T_0 = 2\pi/\omega_0$, and *oscillates* between $x = \pm a$. Here, a is called the *amplitude* of the motion. The parameter ϕ_0, known as the *phase angle*, simply shifts the pattern of motion backward and forward in time. Figure 2.4 shows some examples of simple harmonic motion. Here, $\phi_0 = 0$, $+\pi/4$, and $-\pi/4$ correspond to the solid, short-dashed, and long-dashed curves, respectively.

Note that the frequency, ω_0—and, hence, the period, T_0—of simple harmonic motion is determined by the parameters appearing in the simple harmonic equation, (2.61). However, the amplitude, a, and the phase angle, ϕ_0, are the two integration constants of this second-order ordinary differential equation, and are, thus, determined by the initial conditions: *i.e.*, by the particle's initial displacement and velocity.

Now, from Equations (2.45) and (2.60), the potential energy of our particle at position x is approximately

$$U(x) \simeq \frac{1}{2}\,m\,\omega_0^2\,x^2. \tag{2.63}$$

Hence, the total energy is written

$$E = K + U = \frac{1}{2}\,m\left(\frac{dx}{dt}\right)^2 + \frac{1}{2}\,m\,\omega_0^2\,x^2, \tag{2.64}$$

giving

$$E = \frac{1}{2}\,m\,\omega_0^2\,a^2\,\cos^2(\omega_0\,t - \phi_0) + \frac{1}{2}\,m\,\omega_0^2\,a^2\,\sin^2(\omega_0\,t - \phi_0) = \frac{1}{2}\,m\,\omega_0^2\,a^2, \tag{2.65}$$

where use has been made of Equation (2.62), and the trigonometric identity $\cos^2\theta + \sin^2\theta \equiv 1$. Note that the total energy is *constant* in time, as is to be expected for a conservative system, and is proportional to the *amplitude squared* of the motion.

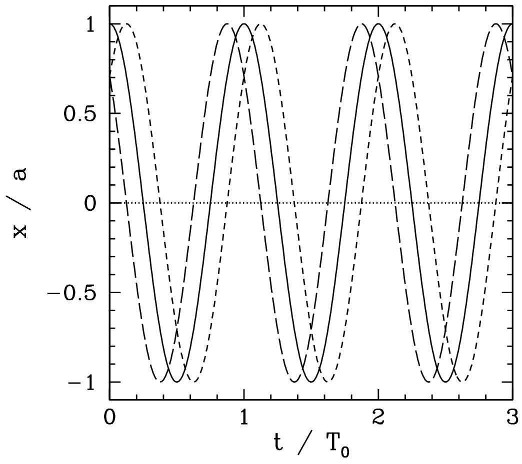


Figure 2.4: *Simple harmonic motion.*

Consider the motion of a point particle of mass m which is slightly displaced from a *unstable* equilibrium point at $x = 0$. The fact that the equilibrium is unstable implies that

$$f(0) = 0, \tag{2.66}$$

and

$$\frac{df(0)}{dx} > 0. \tag{2.67}$$

Now, as long as $|x|$ remains small, our particle's equation of motion takes the approximate form

$$m \frac{d^2x}{dt^2} \simeq f(0) + \frac{df(0)}{dx}\, x, \tag{2.68}$$

which reduces to

$$\frac{d^2x}{dt^2} \simeq k^2\, x, \tag{2.69}$$

where $df(0)/dx = m\, k^2$. The most general solution to the above equation is

$$x(t) = A\, e^{k t} + B\, e^{-k t}, \tag{2.70}$$

where A and B are arbitrary constants. Thus, unless A is *exactly* zero, we conclude that the particle's displacement from the unstable equilibrium point grows *exponentially* in time.

2.9 Two-Body Problem

An isolated dynamical system consisting of two freely moving point masses exerting forces on one another is conventionally termed a *two-body problem*. Suppose that the first object is of mass m_1, and is located at position vector $\mathbf{r}_1$. Likewise, the second object is of mass

m_2, and is located at position vector $\mathbf{r}_2$. Let the first object exert a force $\mathbf{f}_{21}$ on the second. By Newton's third law, the second object exerts an equal and opposite force, $\mathbf{f}_{12} = -\mathbf{f}_{21}$, on the first. Suppose that there are no other forces in the problem. The equations of motion of our two objects are thus

$$m_1 \frac{d^2\mathbf{r}_1}{dt^2} = -\mathbf{f}, \tag{2.71}$$

$$m_2 \frac{d^2\mathbf{r}_2}{dt^2} = \mathbf{f}, \tag{2.72}$$

where $\mathbf{f} = \mathbf{f}_{21}$.

Now, the center of mass of our system is located at

$$\mathbf{r}_{cm} = \frac{m_1\,\mathbf{r}_1 + m_2\,\mathbf{r}_2}{m_1 + m_2}. \tag{2.73}$$

Hence, we can write

$$\mathbf{r}_1 = \mathbf{r}_{cm} - \frac{m_2}{m_1 + m_2}\,\mathbf{r}, \tag{2.74}$$

$$\mathbf{r}_2 = \mathbf{r}_{cm} + \frac{m_1}{m_1 + m_2}\,\mathbf{r}, \tag{2.75}$$

where $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Substituting the above two equations into Equations (2.71) and (2.72), and making use of the fact that the center of mass of an isolated system *does not accelerate* (see Section 2.5), we find that both equations yield

$$\mu\,\frac{d^2\mathbf{r}}{dt^2} = \mathbf{f}, \tag{2.76}$$

where

$$\mu = \frac{m_1\,m_2}{m_1 + m_2} \tag{2.77}$$

is called the *reduced mass*. Hence, we have effectively converted our original two-body problem into an equivalent one-body problem. In the equivalent problem, the force $\mathbf{f}$ is the *same* as that acting on both objects in the original problem (modulo a minus sign). However, the mass, μ, is *different*, and is less than either of m_1 or m_2 (which is why it is called the "reduced" mass). We conclude that the dynamics of an isolated system consisting of two interacting point objects can always be reduced to that of an equivalent system consisting of a single point object moving in a fixed potential.

2.10 Exercises

2.1. Consider a function of many variables $f(x_1, x_2, \cdots, x_n)$. Such a function which satisfies

$$f(t\,x_1, t\,x_2, \cdots, t\,x_n) = t^a\,f(x_1, x_2, \cdots, x_n)$$

for all $t > 0$, and all values of the x_i, is termed a *homogenous function of degree* a. Prove the following theorem regarding homogeneous functions:

$$\sum_{i=1,n} x_i \frac{\partial f}{\partial x_i} = a\, f$$

2.2. Consider an isolated system of N point objects interacting via attractive central forces. Let the mass and position vector of the ith object be m_i and $\mathbf{r}_i$, respectively. Suppose that magnitude of the force exerted on object i by object j is $k_i\, k_j\, |\mathbf{r}_i - \mathbf{r}_j|^{-n}$. Here, the k_i measure some constant physical property of the particles (*e.g.*, their electric charges). Write an expression for the total potential energy U of the system. Is this a homogenous function? If so, what is its degree? Write the equation of motion of the ith particle. Use the mathematical theorem from the previous exercise to demonstrate that

$$\frac{1}{2}\frac{d^2 I}{dt^2} = 2\,K + (n-1)\,U,$$

where $I = \sum_{i=1,N} m_i\, r_i^2$, and K is the kinetic energy. This result is known as the *virial theorem*. Demonstrate that there are no bound steady-state equilibria for the system (*i.e.*, states in which the global system parameters do not evolve in time) when $n \geq 3$.

2.3. A particle of mass m is constrained to move in one dimension such that its instantaneous displacement is x. The particle is released at rest from $x = b$, and is subject to a force of the form $f(x) = -k\, x^{-2}$. Show that the time required for the particle to reach the origin is

$$\pi \left(\frac{m\, b^3}{8\, k}\right)^{1/2}.$$

2.4. A body of uniform cross-sectional area A and mass density ρ floats in a liquid of density ρ_0 (where $\rho < \rho_0$), and at equilibrium displaces a volume V. Show that the period of small oscillations about the equilibrium position is

$$T = 2\pi \sqrt{\frac{V}{g\, A}}.$$

2.5. Using the notation of Section 2.9, show that the total momentum and angular momentum of a two-body system take the form

$$\mathbf{P} = M\, \mathbf{r}_{cm},$$

$$\mathbf{L} = M\, \mathbf{r}_{cm} \times \dot{\mathbf{r}}_{cm} + \mu\, \mathbf{r} \times \dot{\mathbf{r}},$$

respectively, where $M = m_1 + m_2$.

3 Newtonian Gravity

3.1 Introduction

Classical gravity, which is invariably the dominant force in celestial dynamical systems, was first correctly described in Newton's Principia. According to Newton, any two point objects exert a gravitational force of attraction on one another. This force points along the line joining the objects, is directly proportional to the product of their masses, and inversely proportional to the square of the distance between them. Consider two point objects of mass m_1 and m_2 that are located at position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. The force $\mathbf{f}_{12}$ that mass m_2 exerts on mass m_1 is written

$$\mathbf{f}_{12} = G\, m_1\, m_2\, \frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_2 - \mathbf{r}_1|^3}. \tag{3.1}$$

The force $\mathbf{f}_{21}$ that mass m_1 exerts on mass m_2 is equal and opposite: *i.e.*, $\mathbf{f}_{21} = -\mathbf{f}_{12}$. Here, the constant of proportionality, G, is called the *universal gravitational constant*, and takes the value

$$G = 6.67300 \times 10^{-11}\, \mathrm{m}^3\, \mathrm{kg}^{-1}\, \mathrm{s}^{-2}. \tag{3.2}$$

Incidentally, there is something rather curious about Equation (3.1). According to this law, the gravitational force acting on a given object is directly proportional to that object's inertial mass. But why should inertia be related to the force of gravity? After all, inertia measures the reluctance of a given body to deviate from its preferred state of uniform motion in a straight-line, in response to some external force. What has this got to do with gravitational attraction? This question perplexed physicists for many years, and was only answered when Albert Einstein published his general theory of relativity in 1916. According to Einstein, inertial mass acts as a sort of gravitational charge since it impossible to distinguish an acceleration produced by a gravitational field from an apparent acceleration generated by observing in a non-inertial reference frame. The assumption that these two types of acceleration are indistinguishable leads directly to all of the strange predictions of general relativity: *e.g.*, clocks in different gravitational potentials run at different rates, mass bends space, *etc.*

3.2 Gravitational Potential

Consider two point masses, m and m', located at position vectors $\mathbf{r}$ and $\mathbf{r}'$, respectively. According to the above analysis, the acceleration $\mathbf{g}$ of mass m due to the gravitational force exerted on it by mass m' takes the form

$$\mathbf{g} = G\, m'\, \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3}. \tag{3.3}$$

Now, the x-component of this acceleration is written

$$g_x = G\,m'\,\frac{(x'-x)}{[(x'-x)^2+(y'-y)^2+(z'-z)^2]^{3/2}}, \tag{3.4}$$

where $\mathbf{r}=(x,\,y,\,z)$ and $\mathbf{r}'=(x',\,y',\,z')$. However, as is easily demonstrated,

$$\frac{(x'-x)}{[(x'-x)^2+(y'-y)^2+(z'-z)^2]^{3/2}} \equiv \frac{\partial}{\partial x}\left(\frac{1}{[(x'-x)^2+(y'-y)^2+(z'-z)^2]^{1/2}}\right). \tag{3.5}$$

Hence,

$$g_x = G\,m'\,\frac{\partial}{\partial x}\left(\frac{1}{|\mathbf{r}'-\mathbf{r}|}\right), \tag{3.6}$$

with analogous expressions for g_y and g_z. It follows that

$$\mathbf{g}(\mathbf{r}) = -\nabla\Phi, \tag{3.7}$$

where

$$\Phi(\mathbf{r}) = -\frac{G\,m'}{|\mathbf{r}'-\mathbf{r}|} \tag{3.8}$$

is termed the *gravitational potential*. Of course, we can only write $\mathbf{g}$ in the form (3.7) because gravity is a *conservative* force—see Section 2.4.

Now, it is well-known that gravity is a *superposable* force. In other words, the gravitational force exerted on some point mass by a collection of other point masses is simply the vector sum of the forces exerted on the former mass by each of the latter masses taken in isolation. It follows that the gravitational potential generated by a collection of point masses at a certain location in space is the sum of the potentials generated at that location by each point mass taken in isolation. Hence, using Equation (3.8), if there are N point masses, m_i (for $i=1,N$), located at position vectors $\mathbf{r}_i$, then the gravitational potential generated at position vector $\mathbf{r}$ is simply

$$\Phi(\mathbf{r}) = -G\sum_{i=1,N}\frac{m_i}{|\mathbf{r}_i-\mathbf{r}|}. \tag{3.9}$$

Suppose, finally, that, instead of having a collection of point masses, we have a *continuous* mass distribution. In other words, let the mass at position vector $\mathbf{r}'$ be $\rho(\mathbf{r}')\,d^3\mathbf{r}'$, where $\rho(\mathbf{r}')$ is the local mass density, and $d^3\mathbf{r}'$ a volume element. Summing over all space, and taking the limit $d^3\mathbf{r}'\rightarrow 0$, Equation (3.9) yields

$$\Phi(\mathbf{r}) = -G\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}\,d^3\mathbf{r}', \tag{3.10}$$

where the integral is taken over all space. This is the general expression for the gravitational potential, $\Phi(\mathbf{r})$, generated by a continuous mass distribution, $\rho(\mathbf{r})$.

3.3 Gravitational Potential Energy

Consider a collection of N point masses m_i located at position vectors $\mathbf{r}_i$, respectively (where i runs from 1 to N). What is the gravitational potential energy stored in such a collection? In other words, how much work would we have to do in order to assemble the masses, starting from an initial state in which they are all at rest and very widely separated?

We have seen that a gravitational field can be expressed in terms of a gravitational potential: *i.e.*,

$$\mathbf{g}(\mathbf{r}) = -\nabla\Phi. \tag{3.11}$$

We also know that the gravitational force acting on a mass m located at position $\mathbf{r}$ is written

$$\mathbf{f} = m\,\mathbf{g}(\mathbf{r}). \tag{3.12}$$

The work *we* would have to do against the gravitational force in order to *slowly* move the mass from point P to point Q is simply

$$U = -\int_P^Q \mathbf{f}\cdot d\mathbf{r} = -m\int_P^Q \mathbf{g}\cdot d\mathbf{r} = m\int_P^Q \nabla\Phi\cdot d\mathbf{r} = m\,[\Phi(Q)-\Phi(P)]\,. \tag{3.13}$$

The negative sign in the above expression comes about because we would have to exert a force $-\mathbf{f}$ on the mass, in order to counteract the force exerted by the gravitational field. Recall, finally, that the gravitational potential generated by a point mass m' located at position $\mathbf{r}'$ is

$$\Phi(\mathbf{r}) = -\frac{G\,m'}{|\mathbf{r}'-\mathbf{r}|}. \tag{3.14}$$

Let us build up our collection of masses one by one. It takes no work to bring the first mass from infinity, since there is no gravitational field to fight against. Let us clamp this mass in position at $\mathbf{r}_1$. In order to bring the second mass into position at $\mathbf{r}_2$, we have to do work against the gravitational field generated by the first masss. According to Equations (3.13) and Equations (3.14), this work is given by

$$U_2 = -\frac{G\,m_2\,m_1}{|\mathbf{r}_1-\mathbf{r}_2|}. \tag{3.15}$$

Let us now bring the third mass into position. Since gravitational fields and gravitational potentials are superposable, the work done whilst moving the third mass from infinity to $\mathbf{r}_3$ is simply the sum of the works done against the gravitational fields generated by charges 1 and 2 taken in isolation:

$$U_3 = -\frac{G\,m_3\,m_1}{|\mathbf{r}_1-\mathbf{r}_3|} - \frac{G\,m_3\,m_2}{|\mathbf{r}_2-\mathbf{r}_3|}. \tag{3.16}$$

Thus, the total work done in assembling the arrangement of three masses is given by

$$U = -\frac{G\,m_2\,m_1}{|\mathbf{r}_1-\mathbf{r}_2|} - \frac{G\,m_3\,m_1}{|\mathbf{r}_1-\mathbf{r}_3|} - \frac{G\,m_3\,m_2}{|\mathbf{r}_2-\mathbf{r}_3|}. \tag{3.17}$$

This result can easily be generalized to an arrangement of N point masses:

$$U = -\sum_{i=1}^{N}\sum_{j<i}^{N}\frac{G\,m_i\,m_j}{|\mathbf{r}_j - \mathbf{r}_i|}. \tag{3.18}$$

The restriction that j must be less than i makes the above summation rather messy. If we were to sum without restriction (other than $j \neq i$) then each pair of masses would be counted twice. It is convenient to do just this, and then to divide the result by two. Thus, we obtain

$$U = -\frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{G\,m_i\,m_j}{|\mathbf{r}_j - \mathbf{r}_i|}. \tag{3.19}$$

This is the *potential energy* of an arrangement of point masses. We can think of this quantity as the work required to bring the masses from infinity and assemble them in the required formation. Alternatively, it is the kinetic energy that would be released if the assemblage were dissolved, and the masses returned to infinity.

Equation (3.19) can be written

$$W = \frac{1}{2}\sum_{i=1}^{N} m_i\,\Phi_i, \tag{3.20}$$

where

$$\Phi_i = -G\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{m_j}{|\mathbf{r}_j - \mathbf{r}_i|} \tag{3.21}$$

is the gravitational potential experienced by the ith mass due to the other masses in the distribution. For the case of a continuous mass distribution, we can generalize the above result to give

$$U = \frac{1}{2}\int \rho(\mathbf{r})\,\Phi(\mathbf{r})\,d^3\mathbf{r}, \tag{3.22}$$

where

$$\Phi(\mathbf{r}) = -G\int\frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}\,d^3\mathbf{r}' \tag{3.23}$$

is the familiar gravitational potential generated by a continuous mass distribution of mass density $\rho(\mathbf{r})$.

3.4 Axially Symmetric Mass Distributions

At this point, it is convenient to adopt standard spherical coordinates, (r, θ, ϕ), aligned along the z-axis. These coordinates are related to regular Cartesian coordinates as follows

(see Section A.17):

$$x = r \sin\theta \cos\phi, \tag{3.24}$$

$$y = r \sin\theta \sin\phi, \tag{3.25}$$

$$z = r \cos\theta. \tag{3.26}$$

Consider an *axially symmetric* mass distribution: *i.e.*, a $\rho(\mathbf{r})$ which is *independent* of the azimuthal angle, ϕ. We would expect such a mass distribution to generated an axially symmetric gravitational potential, $\Phi(r,\theta)$. Hence, without loss of generality, we can set $\phi = 0$ when evaluating $\Phi(\mathbf{r})$ from Equation (3.23). In fact, given that $d^3\mathbf{r}' = r'^2 \sin\theta'\, dr'\, d\theta'\, d\phi'$ in spherical coordinates, this equation yields

$$\Phi(r,\theta) = -G \int_0^{\infty}\int_0^{\pi}\int_0^{2\pi} \frac{r'^2\,\rho(r',\theta')\,\sin\theta'}{|\mathbf{r}-\mathbf{r}'|}\,dr'\,d\theta'\,d\phi', \tag{3.27}$$

with the right-hand side evaluated at $\phi = 0$. However, since $\rho(r',\theta')$ is independent of ϕ', the above equation can also be written

$$\Phi(r,\theta) = -2\pi\, G \int_0^{\infty}\int_0^{\pi} r'^2\,\rho(r',\theta')\,\sin\theta'\,\langle|\mathbf{r}-\mathbf{r}'|^{-1}\rangle\,dr'\,d\theta', \tag{3.28}$$

where $\langle\cdots\rangle \equiv \oint(\cdots)\,d\phi'/2\pi$ denotes an average over the azimuthal angle, ϕ'.

Now,

$$|\mathbf{r}'-\mathbf{r}|^{-1} = (r^2 - 2\,\mathbf{r}\cdot\mathbf{r}' + r'^2)^{-1/2}, \tag{3.29}$$

and

$$\mathbf{r}\cdot\mathbf{r}' = r\,r'\,F, \tag{3.30}$$

where (at $\phi = 0$)

$$F = \sin\theta\,\sin\theta'\,\cos\phi' + \cos\theta\,\cos\theta'. \tag{3.31}$$

Hence,

$$|\mathbf{r}'-\mathbf{r}|^{-1} = (r^2 - 2\,r\,r'\,F + r'^2)^{-1/2}. \tag{3.32}$$

Suppose that $r > r'$. In this case, we can expand $|\mathbf{r}'-\mathbf{r}|^{-1}$ as a convergent power series in r'/r, to give

$$|\mathbf{r}'-\mathbf{r}|^{-1} = \frac{1}{r}\left[1 + \left(\frac{r'}{r}\right)F + \frac{1}{2}\left(\frac{r'}{r}\right)^2(3\,F^2-1) + \mathcal{O}\left(\frac{r'}{r}\right)^3\right]. \tag{3.33}$$

Let us now average this expression over the azimuthal angle, ϕ'. Since $\langle 1\rangle = 1$, $\langle\cos\phi'\rangle = 0$, and $\langle\cos^2\phi'\rangle = 1/2$, it is easily seen that

$$\langle F\rangle = \cos\theta\,\cos\theta', \tag{3.34}$$

$$\begin{aligned}\langle F^2\rangle &= \frac{1}{2}\sin^2\theta\,\sin^2\theta' + \cos^2\theta\,\cos^2\theta' \\ &= \frac{1}{3} + \frac{2}{3}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\left(\frac{3}{2}\cos^2\theta' - \frac{1}{2}\right).\end{aligned} \tag{3.35}$$

Hence,

$$\left\langle |\mathbf{r}' - \mathbf{r}|^{-1} \right\rangle = \frac{1}{r}\left[1 + \left(\frac{r'}{r}\right)\cos\theta\,\cos\theta' + \left(\frac{r'}{r}\right)^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\left(\frac{3}{2}\cos^2\theta' - \frac{1}{2}\right) + \mathcal{O}\left(\frac{r'}{r}\right)^3\right]. \tag{3.36}$$

Now, the well-known *Legendre polynomials,* $P_n(x)$, are defined

$$P_n(x) = \frac{1}{2^n\,n!}\,\frac{d^n}{dx^n}\left[(x^2-1)^n\right], \tag{3.37}$$

for $n = 0, \infty$. It follows that

$$P_0(x) = 1, \tag{3.38}$$

$$P_1(x) = x, \tag{3.39}$$

$$P_2(x) = \frac{1}{2}\,(3\,x^2 - 1), \tag{3.40}$$

etc. The Legendre polynomials are *mutually orthogonal*: *i.e.*,

$$\int_{-1}^{1} P_n(x)\,P_m(x)\,dx = \int_0^{\pi} P_n(\cos\theta)\,P_m(\cos\theta)\,\sin\theta\,d\theta = \frac{\delta_{nm}}{n+1/2}. \tag{3.41}$$

Here, δ_{nm} is 1 if $n = m$, and 0 otherwise. The Legendre polynomials also form a *complete set*: *i.e.*, any well-behaved function of x can be represented as a weighted sum of the $P_n(x)$. Likewise, any well-behaved (even) function of θ can be represented as a weighted sum of the $P_n(\cos\theta)$.

A comparison of Equation (3.36) and Equations (3.38)–(3.40) makes it reasonably clear that, when $r > r'$, the complete expansion of $\langle |\mathbf{r}' - \mathbf{r}|^{-1}\rangle$ is

$$\left\langle |\mathbf{r}' - \mathbf{r}|^{-1} \right\rangle = \frac{1}{r}\sum_{n=0,\infty}\left(\frac{r'}{r}\right)^n P_n(\cos\theta)\,P_n(\cos\theta'). \tag{3.42}$$

Similarly, when $r < r'$, we can expand in powers of r/r' to obtain

$$\left\langle |\mathbf{r}' - \mathbf{r}|^{-1} \right\rangle = \frac{1}{r'}\sum_{n=0,\infty}\left(\frac{r}{r'}\right)^n P_n(\cos\theta)\,P_n(\cos\theta'). \tag{3.43}$$

It follows from Equations (3.28), (3.42), and (3.43) that

$$\Phi(r,\theta) = \sum_{n=0,\infty} \Phi_n(r)\,P_n(\cos\theta), \tag{3.44}$$

where

$$\Phi_n(r) = -\frac{2\pi\, G}{r^{n+1}}\int_0^r\int_0^\pi r'^{\,n+2}\,\rho(r',\theta')\,P_n(\cos\theta')\,\sin\theta'\,dr'\,d\theta'$$
$$-2\pi\,G\,r^n\int_r^\infty\int_0^\pi r'^{\,1-n}\,\rho(r',\theta')\,P_n(\cos\theta')\,\sin\theta'\,dr'\,d\theta'. \tag{3.45}$$

Now, given that the $P_n(\cos\theta)$ form a complete set, we can always write

$$\rho(r,\theta) = \sum_{n=0,\infty}\rho_n(r)\,P_n(\cos\theta). \tag{3.46}$$

This expression can be inverted, with the aid of Equation (3.41), to give

$$\rho_n(r) = (n+1/2)\int_0^\pi \rho(r,\theta)\,P_n(\cos\theta)\,\sin\theta\,d\theta. \tag{3.47}$$

Hence, Equation (3.45) reduces to

$$\Phi_n(r) = -\frac{2\pi\,G}{(n+1/2)\,r^{n+1}}\int_0^r r'^{\,n+2}\,\rho_n(r')\,dr' - \frac{2\,\pi\,G\,r^n}{n+1/2}\int_r^\infty r'^{\,1-n}\,\rho_n(r')\,dr'. \tag{3.48}$$

Thus, we now have a general expression for the gravitational potential, $\Phi(r,\theta)$, generated by any axially symmetric mass distribution, $\rho(r,\theta)$.

3.5 Potential Due to a Uniform Sphere

Let us calculate the gravitational potential generated by a sphere of uniform mass density γ and radius a. Expressing $\rho(r,\theta)$ in the form (3.46), it is clear that

$$\rho_0(r) = \left\{\begin{array}{ll}\gamma & \text{for } r\leq a\\ 0 & \text{for } r>a\end{array}\right., \tag{3.49}$$

with $\rho_n(r) = 0$ for $n>0$. Thus, from (3.48),

$$\Phi_0(r) = -\frac{4\pi\,G\,\gamma}{r}\int_0^r r'^2\,dr' - 4\pi\,G\,\gamma\int_r^a r'\,dr' \tag{3.50}$$

for $r\leq a$, and

$$\Phi_0(r) = -\frac{4\pi\,G\,\gamma}{r}\int_0^a r'^2\,dr' \tag{3.51}$$

for $r>a$, with $\Phi_n(r) = 0$ for $n>0$. Hence,

$$\Phi(r) = -\frac{2\pi\,G\,\gamma}{3}\,(3\,a^2-r^2) = -G\,M\,\frac{(3\,a^2-r^2)}{2\,a^3} \tag{3.52}$$

for $r \leq a$, and

$$\Phi(r) = -\frac{4\pi\,G\,\gamma}{3}\,\frac{a^3}{r} = -\frac{G\,M}{r} \tag{3.53}$$

for $r > a$. Here, $M = (4\pi/3)\,a^3\,\gamma$ is the total mass of the sphere.

According to Equation (3.53), the gravitational potential *outside* a uniform sphere of mass M is the same as that generated by a point mass M located at the sphere's center. It turns out that this is a general result for *any* finite spherically symmetric mass distribution. Indeed, from the above analysis, it is clear that $\rho(r) = \rho_0(r)$ and $\Phi(r) = \Phi_0(r)$ for a spherically symmetric mass distribution. Suppose that the mass distribution extends out to $r = a$. It immediately follows, from Equation (3.48), that

$$\Phi(r) = -\frac{G}{r}\int_0^a 4\pi\,r'^2\,\rho(r')\,dr' = -\frac{G\,M}{r} \tag{3.54}$$

for $r > a$, where M is the total mass of the distribution. Now the center of mass of a spherically symmetric mass distribution lies at the center of the distribution. Moreover, the translational motion of the center of mass is analogous to that of a point particle, whose mass is equal to that of the whole distribution, moving under the action of the net external force (see Section 2.6). We, thus, conclude that Newton's laws of motion, in their primitive form, apply not only to point masses, but also to the translational motions of extended spherically symmetric mass distributions interacting via gravity (*e.g.*, the Sun and the Planets).

3.6 Potential Outside a Uniform Spheroid

Let us now calculate the gravitational potential generated outside a *spheroid* of uniform mass density γ and mean radius a. A spheroid is the solid body produced by rotating an ellipse about a major or minor axis. Let the axis of rotation coincide with the z-axis, and let the outer boundary of the spheroid satisfy

$$r = a_\theta(\theta) = a\left[1 - \frac{2}{3}\,\epsilon\,P_2(\cos\theta)\right], \tag{3.55}$$

where ϵ is the termed the *ellipticity*. Here, we are assuming that $|\epsilon| \ll 1$, so that the spheroid is very close to being a sphere. If $\epsilon > 0$ then the spheroid is slightly squashed along its symmetry axis, and is termed *oblate*. Likewise, if $\epsilon < 0$ then the spheroid is slightly elongated along its axis, and is termed *prolate*—see Figure 3.1. Of course, if $\epsilon = 0$ then the spheroid reduces to a sphere.

Now, according to Equation (3.44) and (3.45), the gravitational potential generated *outside* an axially symmetric mass distribution can be written

$$\Phi(r,\theta) = \sum_{n=0,\infty} J_n\,\frac{P_n(\cos\theta)}{r^{n+1}}, \tag{3.56}$$

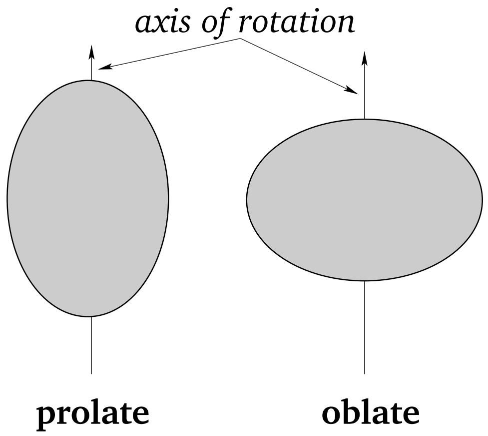


Figure 3.1: *Prolate and oblate spheroids.*

where

$$J_n = -2\pi\, G \int\!\!\int r^{\,2+n}\, \rho(r,\theta)\, P_n(\cos\theta)\, \sin\theta\, dr\, d\theta. \tag{3.57}$$

Here, the integral is taken over the whole cross-section of the distribution in r-θ space.

It follows that for a uniform spheroid

$$J_n = -2\pi\, G\, \gamma \int_0^{\pi} P_n(\cos\theta) \int_0^{a_\theta(\theta)} r^{\,2+n}\, dr\, \sin\theta\, d\theta \tag{3.58}$$

Hence,

$$J_n = -\frac{2\pi\, G\, \gamma}{(3+n)} \int_0^{\pi} P_n(\cos\theta)\, a_\theta^{3+n}(\theta)\, \sin\theta\, d\theta, \tag{3.59}$$

giving

$$J_n \simeq -\frac{2\pi\, G\, \gamma\, a^{3+n}}{(3+n)} \int_0^{\pi} P_n(\cos\theta) \left[P_0(\cos\theta) - \frac{2}{3}\,(3+n)\, \epsilon\, P_2(\cos\theta)\right] \sin\theta\, d\theta, \tag{3.60}$$

to first-order in ϵ. It is thus clear, from Equation (3.41), that, to first-order in ϵ, the only non-zero J_n are

$$J_0 = -\frac{4\pi\, G\, \gamma\, a^3}{3} = -G\, M, \tag{3.61}$$

$$J_2 = \frac{8\pi\, G\, \gamma\, a^5\, \epsilon}{15} = \frac{2}{5}\, G\, M\, a^2\, \epsilon, \tag{3.62}$$

where $M = (4\pi/3)\, a^3\, \gamma$ is the total mass.

Thus, the gravitational potential outside a uniform spheroid of total mass M, mean radius a, and ellipticity ϵ, is

$$\Phi(r,\theta) = -\frac{G\, M}{r} + \frac{2}{5}\frac{G\, M\, a^2}{r^3}\, \epsilon\, P_2(\cos\theta) + \mathcal{O}(\epsilon^2). \tag{3.63}$$

In particular, the gravitational potential on the surface of the spheroid is

$$\Phi(a_\theta, \theta) = -\frac{G\,M}{a_\theta} + \frac{2}{5}\frac{G\,M\,a^2}{a_\theta^3}\,\epsilon\,P_2(\cos\theta) + \mathcal{O}(\epsilon^2), \tag{3.64}$$

which yields

$$\Phi(a_\theta, \theta) \simeq -\frac{G\,M}{a}\left[1 + \frac{4}{15}\,\epsilon\,P_2(\cos\theta) + \mathcal{O}(\epsilon^2)\right], \tag{3.65}$$

where use has been made of Equation (3.55).

Consider a self-gravitating spheroid of mass M, mean radius a, and ellipticity ϵ: *e.g.*, a star, or a planet. Assuming, for the sake of simplicity, that the spheroid is composed of uniform density incompressible fluid, the gravitational potential on its surface is given by Equation (3.65). However, the condition for an equilibrium state is that the potential be *constant* over the surface. If this is not the case then there will be gravitational forces acting *tangential* to the surface. Such forces cannot be balanced by internal pressure, which only acts *normal* to the surface. Hence, from (3.65), it is clear that the condition for equilibrium is $\epsilon = 0$. In other words, the equilibrium configuration of a self-gravitating mass is a *sphere*. Deviations from this configuration can only be caused by forces in addition to self-gravity and internal pressure: *e.g.*, internal tensile forces, centrifugal forces due to rotation, or tidal forces due to orbiting masses (see Chapter 6).

We can estimate how small a rocky asteroid or moon needs to be before its internal tensile strength is sufficient to allow it to retain a significantly non-spherical shape. The typical density of rock is $\gamma \sim 3\times 10^3\,\mathrm{kg\,m^{-3}}$. Moreover, the critical compressional stress at which rock ceases to act like a rigid material, and instead deforms and flows like a liquid, is $P_c \sim 10^7\,\mathrm{N\,m^{-2}}$. We must compare this critical stress with the pressure at the center of the asteroid or moon. Assuming, for the sake of simplicity, that the asteroid or moon is spherical, of radius a, and of uniform density γ, the central pressure is

$$P_0 = \int_0^a \gamma\,g(r)\,dr, \tag{3.66}$$

where $g(r) = (4\pi/3)\,G\,\gamma\,r$ is the gravitational acceleration at radius r [see Equation (3.52)]. The above result is a simple generalization of the well-known formula $\rho\,g\,h$ for the pressure a depth h below the surface of a fluid. It follows that

$$P_0 = \frac{2\pi}{3}\,G\,\gamma^2\,a^2. \tag{3.67}$$

Now, if $P_0 \ll P_c$ then the internal pressure in the asteroid or moon is not sufficiently high to cause its constituent rock to deform like a liquid. Such an asteroid or moon can therefore retain a significantly non-spherical shape. On the other hand, if $P_0 \gg P_c$ then the internal pressure is large enough to render the asteroid or moon fluid-like. Such a body cannot withstand the tendency of self-gravity to make it adopt a spherical shape. The condition $P_0 \ll P_c$ is equivalent to $a \ll a_c$, where

$$a_c = \left(\frac{3}{2\pi}\,\frac{P_c}{G\,\gamma^2}\right)^{1/2} \sim 100\,\mathrm{km}. \tag{3.68}$$

It follows that only a rocky asteroid or moon whose radius is significantly less than about 100 km (*e.g.*, the two moons of Mars, Phobos and Deimos) can retain a highly non-spherical shape. On the other hand, an asteroid or moon whose radius is significantly greater than about 100 km (*e.g.*, the asteroid Ceres, and the Earth's moon) is forced to be essentially spherical.

3.7 Potential Due to a Uniform Ring

Consider a uniform ring of mass M and radius r, centered on the origin, and lying in the x-y plane. Let us consider the gravitational potential $\Phi(r)$ generated by such a ring in the x-y plane (which corresponds to $\theta = 90^\circ$). It follows, from Section 3.4, that for $r > a$,

$$\Phi(r) = -\frac{G\,M}{a} \sum_{n=0,\infty} [P_n(0)]^2 \left(\frac{a}{r}\right)^{n+1}. \tag{3.69}$$

However, $P_0(0) = 1$, $P_1(0) = 0$, $P_2(0) = -1/2$, $P_3(0) = 0$, and $P_4(0) = 3/8$. Hence,

$$\Phi(r) = -\frac{G\,M}{r}\left[1 + \frac{1}{4}\left(\frac{a}{r}\right)^2 + \frac{9}{64}\left(\frac{a}{r}\right)^4 + \cdots\right]. \tag{3.70}$$

Likewise, for $r < a$,

$$\Phi(r) = -\frac{G\,M}{a} \sum_{n=0,\infty} [P_n(0)]^2 \left(\frac{r}{a}\right)^{n}, \tag{3.71}$$

giving

$$\Phi(r) = -\frac{G\,M}{a}\left[1 + \frac{1}{4}\left(\frac{r}{a}\right)^2 + \frac{9}{64}\left(\frac{r}{a}\right)^4 + \cdots\right]. \tag{3.72}$$

3.8 Exercises

3.1. Consider an isolated system of N point objects interacting via gravity. Let the mass and position vector of the ith object be m_i and $\mathbf{r}_i$, respectively. What is the vector equation of motion of the ith object? Write expressions for the total kinetic energy, K, and potential energy, U, of the system. Demonstrate from the equations of motion that $K+U$ is a conserved quantity.

3.2. A particle is projected vertically upward from the Earth's surface with a velocity which would, if gravity were uniform, carry it to a height h. Show that if the variation of gravity with height is allowed for, but the resistance of air is neglected, then the height reached will be greater by $h^2/(R-h)$, where R is the Earth's radius.

3.3. A particle is projected vertically upward from the Earth's surface with a velocity just sufficient for it to reach infinity (neglecting air resistance). Prove that the time needed to reach a height h is

$$\frac{1}{3}\left(\frac{2\,R}{g}\right)^{1/2}\left[\left(1+\frac{h}{R}\right)^{3/2} - 1\right].$$

where R is the Earth's radius, and g its surface gravitational acceleration.

3.4. Demonstrate that if a narrow shaft were drilled though the center of a uniform self-gravitating sphere then a test mass moving in this shaft executes simple harmonic motion about the center of the sphere with period

$$T = 2\pi\sqrt{\frac{a}{g}},$$

where a is the radius of the sphere, and g the gravitational acceleration at its surface.

3.5. Demonstrate that the gravitational potential energy of a spherically symmetric mass distribution of mass density $\rho(r)$ that extends out to $r = a$ can be written

$$U = -16\pi^2\,G\int_0^a r\,\rho(r)\int_0^r r'^2\,\rho(r')\,dr'\,dr.$$

Hence, show that if the mass distribution is such that

$$\rho(r) = \begin{cases} \rho_0\,r^{-\alpha} & r \leq a \\ 0 & r > a \end{cases},$$

where $\alpha < 5/2$, then

$$U = -\frac{(3-\alpha)}{(5-2\,\alpha)}\,\frac{G\,M^2}{a},$$

where M is the total mass.

3.6. A star can be through of as a spherical system that consists of a very large number of particles interacting via gravity. Show that, for such a system, the virial theorem, introduced in Exercise 2.2, implies that

$$\frac{d^2I}{dt^2} = -2\,U + c,$$

where c is a constant, and the r_i are measured from the geometric center. Hence, deduce that the angular frequency of small amplitude radial pulsations of the star (in which the radial displacement is directly proportional to the radial distance from the center) takes the form

$$\omega = \left(\frac{|U_0|}{I_0}\right)^{1/2},$$

where U_0 and I_0 are the equilibrium values of U and I. Finally, show that if the mass density within the star varies as $r^{-\alpha}$, where r is the radial distance from the geometric center, and where $\alpha < 5/2$, then

$$\omega = \left(\frac{5-\alpha}{5-2\,\alpha}\,\frac{G\,M}{R^3}\right)^{1/2},$$

where M and R are the stellar mass and radius, respectively.

4 Keplerian Orbits

4.1 Introduction

Newtonian mechanics was initially developed in order to account for the motion of the Planets in the Solar System. Let us now examine this problem. Suppose that the Sun, whose mass is M, is located at the origin of our coordinate system. Consider the motion of a general planet, of mass m, that is located at position vector $\mathbf{r}$. Given that the Sun and all of the Planets are approximately spherical, the gravitational force exerted on the planet by the Sun can be written (see Chapter 3)

$$\mathbf{f} = -\frac{G\,M\,m}{r^3}\,\mathbf{r}. \tag{4.1}$$

An equal and opposite force to (4.1) acts on the Sun. However, we shall assume that the Sun is so much more massive than the planet that this force does not cause the Sun's position to shift appreciably. Hence, the Sun will always remain at the origin of our coordinate system. Likewise, we shall neglect the gravitational forces exerted on our chosen planet by the other bodies in the Solar System, compared to the gravitational force exerted on it by the Sun. This is reasonable because the Sun is much more massive (by a factor of, at least, 10^3) than any other body in the Solar System. Thus, according to Equation (4.1), and Newton's second law of motion, the equation of motion of our planet (which can effectively be treated as a point object—see Chapter 3) takes the form

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{G\,M}{r^3}\,\mathbf{r}. \tag{4.2}$$

Note that the planet's mass, m, has canceled out on both sides of the above equation.

4.2 Kepler's Laws

As is well-known, Johannes Kepler was the first astronomer to correctly describe the motion of the Planets in the Solar System (in works published between 1609 and 1619). The motion of the Planets is summed up in three simple laws:

1. The planetary orbits are all ellipses which are confocal with the Sun (*i.e.*, the Sun lies at one of the focii of each ellipse—see Appendix B.1).

2. The radius vectors connecting each planet to the Sun sweep out equal areas in equal time intervals.

3. The squares of the orbital periods of the planets are proportional to the cubes of their orbital major radii.

Let us now see if we can derive Kepler's laws from Equation (4.2).

4.3 Conservation Laws

As we have already seen, gravity is a *conservative force*. Hence, the gravitational force (4.1) can be written (see Section 2.4)

$$\mathbf{f} = -\nabla U, \tag{4.3}$$

where the potential energy, $U(\mathbf{r})$, of our planet in the Sun's gravitational field takes the form (see Section 3.5)

$$U(\mathbf{r}) = -\frac{G\,M\,m}{r}. \tag{4.4}$$

It follows that the *total energy* of our planet is a conserved quantity—see Section 2.4. In other words,

$$\mathcal{E} = \frac{v^2}{2} - \frac{G\,M}{r} \tag{4.5}$$

is constant in time. Here, $\mathcal{E}$ is actually the planet's total energy *per unit mass*, and $\mathbf{v} = d\mathbf{r}/dt$.

Gravity is also a *central force*. Hence, the *angular momentum* of our planet is a conserved quantity—see Section 2.5. In other words,

$$\mathbf{h} = \mathbf{r} \times \mathbf{v}, \tag{4.6}$$

which is actually the planet's angular momentum *per unit mass*, is constant in time. Taking the scalar product of the above equation with $\mathbf{r}$, we obtain

$$\mathbf{h} \cdot \mathbf{r} = 0. \tag{4.7}$$

This is the equation of a *plane* which passes through the origin, and whose normal is parallel to $\mathbf{h}$. Since $\mathbf{h}$ is a constant vector, it always points in the *same* direction. We, therefore, conclude that the motion of our planet is *two-dimensional* in nature: *i.e.*, it is confined to some fixed plane which passes through the origin. Without loss of generality, we can let this plane coincide with the x-y plane.

4.4 Polar Coordinates

We can determine the instantaneous position of our planet in the x-y plane in terms of standard Cartesian coordinates, $(x,\ y)$, or polar coordinates, $(r,\ \theta)$, as illustrated in Figure 4.1. Here, $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$. It is helpful to define two unit vectors, $\mathbf{e}_r \equiv \mathbf{r}/r$ and $\mathbf{e}_\theta \equiv \mathbf{e}_z \times \mathbf{e}_r$, at the instantaneous position of the planet. The first always points radially away from the origin, whereas the second is normal to the first, in the direction of increasing θ. As is easily demonstrated, the Cartesian components of $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are

$$\mathbf{e}_r = (\cos\theta,\ \sin\theta), \tag{4.8}$$

$$\mathbf{e}_\theta = (-\sin\theta,\ \cos\theta), \tag{4.9}$$

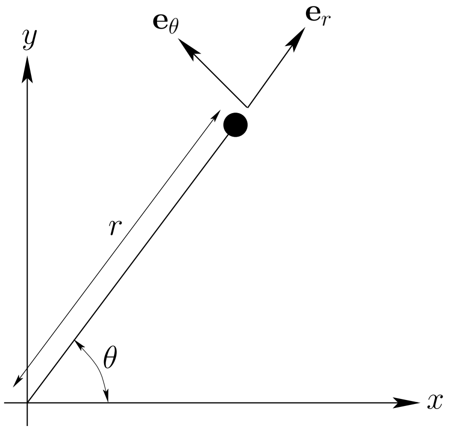


Figure 4.1: *Polar coordinates.*

respectively.

We can write the position vector of our planet as

$$\mathbf{r} = r\,\mathbf{e}_r. \tag{4.10}$$

Thus, the planet's velocity becomes

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \dot{r}\,\mathbf{e}_r + r\,\dot{\mathbf{e}}_r, \tag{4.11}$$

where $\dot{}$ is shorthand for d/dt. Note that $\mathbf{e}_r$ has a non-zero time-derivative (unlike a Cartesian unit vector) because its direction *changes* as the planet moves around. As is easily demonstrated, from differentiating Equation (4.8) with respect to time,

$$\dot{\mathbf{e}}_r = \dot{\theta}\,(-\sin\theta,\,\cos\theta) = \dot{\theta}\;\mathbf{e}_\theta. \tag{4.12}$$

Thus,

$$\mathbf{v} = \dot{r}\;\mathbf{e}_r + r\,\dot{\theta}\;\mathbf{e}_\theta. \tag{4.13}$$

Now, the planet's acceleration is written

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \ddot{r}\,\mathbf{e}_r + \dot{r}\,\dot{\mathbf{e}}_r + (\dot{r}\,\dot{\theta} + r\,\ddot{\theta})\,\mathbf{e}_\theta + r\,\dot{\theta}\;\dot{\mathbf{e}}_\theta. \tag{4.14}$$

Again, $\mathbf{e}_\theta$ has a non-zero time-derivative because its direction *changes* as the planet moves around. Differentiation of Equation (4.9) with respect to time yields

$$\dot{\mathbf{e}}_\theta = \dot{\theta}\,(-\cos\theta,\,-\sin\theta) = -\dot{\theta}\,\mathbf{e}_r. \tag{4.15}$$

Hence,

$$\mathbf{a} = (\ddot{r} - r\,\dot{\theta}^{\,2})\,\mathbf{e}_r + (r\,\ddot{\theta} + 2\,\dot{r}\,\dot{\theta})\,\mathbf{e}_\theta. \tag{4.16}$$

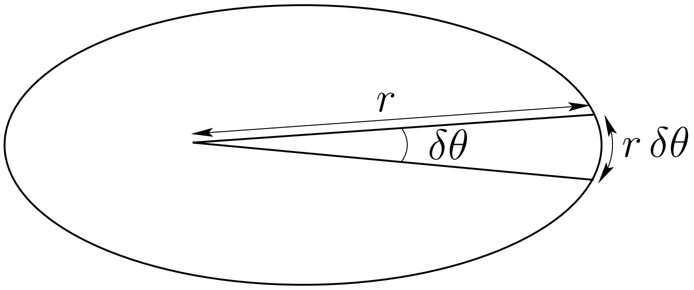


Figure 4.2: *Kepler's second law.*

It follows that the equation of motion of our planet, (4.2), can be written

$$\mathbf{a} = (\ddot{r} - r\,\dot{\theta}^2)\,\mathbf{e}_r + (r\,\ddot{\theta} + 2\,\dot{r}\,\dot{\theta})\,\mathbf{e}_\theta = -\frac{G\,M}{r^2}\,\mathbf{e}_r. \tag{4.17}$$

Since $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are mutually orthogonal, we can separately equate the coefficients of both, in the above equation, to give a *radial equation of motion*,

$$\ddot{r} - r\,\dot{\theta}^2 = -\frac{G\,M}{r^2}, \tag{4.18}$$

and a *tangential equation of motion*,

$$r\,\ddot{\theta} + 2\,\dot{r}\,\dot{\theta} = 0. \tag{4.19}$$

4.5 Kepler's Second Law

Multiplying our planet's tangential equation of motion, (4.19), by r, we obtain

$$r^2\,\ddot{\theta} + 2\,r\,\dot{r}\,\dot{\theta} = 0. \tag{4.20}$$

However, the above equation can be also written

$$\frac{d(r^2\,\dot{\theta})}{dt} = 0, \tag{4.21}$$

which implies that

$$h = r^2\,\dot{\theta} \tag{4.22}$$

is constant in time. It is easily demonstrated that h is the magnitude of the vector $\mathbf{h}$ defined in Equation (4.6). Thus, the fact that h is constant in time is equivalent to the statement that the angular momentum of our planet is a constant of its motion. As we have already mentioned, this is the case because gravity is a central force.

Suppose that the radius vector connecting our planet to the origin (*i.e.*, the Sun) sweeps out an angle $\delta\theta$ between times t and $t+\delta t$—see Figure 4.2. The approximately triangular region swept out by the radius vector has the area

$$\delta A \simeq \frac{1}{2}\,r^2\,\delta\theta, \tag{4.23}$$

since the area of a triangle is half its base ($r\,\delta\theta$) times its height (r). Hence, the rate at which the radius vector sweeps out area is

$$\frac{dA}{dt} = \lim_{\delta t\to 0}\frac{r^2\,\delta\theta}{2\,\delta t} = \frac{r^2}{2}\,\frac{d\theta}{dt} = \frac{h}{2}. \tag{4.24}$$

Thus, the radius vector sweeps out area at a constant rate (since h is constant in time)—this is Kepler's second law. We conclude that Kepler's second law of planetary motion is a direct consequence of *angular momentum conservation*.

4.6 Kepler's First Law

Our planet's radial equation of motion, (4.18), can be combined with Equation (4.22) to give

$$\ddot{r} - \frac{h^2}{r^3} = -\frac{G\,M}{r^2}. \tag{4.25}$$

Suppose that $r = u^{-1}$. It follows that

$$\dot{r} = -\frac{\dot{u}}{u^2} = -r^2\,\frac{du}{d\theta}\,\frac{d\theta}{dt} = -h\,\frac{du}{d\theta}. \tag{4.26}$$

Likewise,

$$\ddot{r} = -h\,\frac{d^2u}{d\theta^2}\,\dot{\theta} = -u^2\,h^2\,\frac{d^2u}{d\theta^2}. \tag{4.27}$$

Hence, Equation (4.25) can be written in the *linear* form

$$\frac{d^2u}{d\theta^2} + u = \frac{G\,M}{h^2}. \tag{4.28}$$

The general solution to the above equation is

$$u(\theta) = \frac{G\,M}{h^2}\,[1 - e\,\cos(\theta-\theta_0)]\,, \tag{4.29}$$

where e and θ_0 are arbitrary constants. Without loss of generality, we can set $\theta_0 = 0$ by rotating our coordinate system about the z-axis. Thus, we obtain

$$r(\theta) = \frac{r_c}{1 - e\,\cos\theta}, \tag{4.30}$$

where

$$r_c = \frac{h^2}{G\,M}. \tag{4.31}$$

We immediately recognize Equation (4.30) as the equation of a conic section that is confocal with the origin (*i.e.*, with the Sun)—see Appendix B.1. Specifically, for $e < 1$, (4.30) is the equation of an *ellipse*. For $e = 1$, (4.30) is the equation of a *parabola*. Finally, for $e > 1$, (4.30) is the equation of a *hyperbola*. However, a planet cannot have a parabolic or a hyperbolic orbit, since such orbits are only appropriate to objects which are ultimately able to escape from the Sun's gravitational field. Thus, the orbit of our planet around the Sun is a confocal ellipse—this is Kepler's first law of planetary motion

4.7 Kepler's Third Law

We have seen that the radius vector connecting our planet to the origin sweeps out area at the constant rate $dA/dt = h/2$ [see Equation (4.24)]. We have also seen that the planetary orbit is an ellipse. Suppose that the major and minor radii of the ellipse are a and b, respectively. It follows that the area of the ellipse is $A = \pi\,a\,b$. Now, we expect the radius vector to sweep out the whole area of the ellipse in a single orbital period, T. Hence,

$$T = \frac{A}{(dA/dt)} = \frac{2\pi\,a\,b}{h}. \tag{4.32}$$

It follows from Equations (B.7), (B.8), and (4.31) that

$$T^2 = \frac{4\pi^2\,a^3}{G\,M}. \tag{4.33}$$

In other words, the *square* of the orbital *period* of our planet is proportional to the *cube* of its orbital *major radius*—this is Kepler's third law.

Note that for an elliptical orbit the closest distance to the Sun—the so-called *perihelion* distance—is [see Equations (B.7) and (4.30)]

$$r_p = \frac{r_c}{1+e} = a\,(1-e). \tag{4.34}$$

Likewise, the furthest distance from the Sun—the so-called *aphelion* distance—is

$$r_a = \frac{r_c}{1-e} = a\,(1+e). \tag{4.35}$$

It follows that the major radius, a, is simply the mean of the perihelion and aphelion distances,

$$a = \frac{r_p + r_a}{2}. \tag{4.36}$$

The parameter

$$e = \frac{r_a - r_p}{r_a + r_p} \tag{4.37}$$

is called the *eccentricity*, and measures the deviation of the orbit from circularity. Thus, $e = 0$ corresponds to a circular orbit, whereas $e \to 1$ corresponds to an infinitely elongated elliptical orbit.

As is easily demonstrated from the above analysis, Kepler laws of planetary motion can be written in the convenient form

$$r = \frac{a\,(1-e^2)}{1-e\,\cos\theta}, \tag{4.38}$$

$$r^2\,\dot{\theta} = (1-e^2)^{1/2}\,n\,a^2, \tag{4.39}$$

$$G\,M = n^2\,a^3, \tag{4.40}$$

where a is the mean orbital radius (*i.e.*, the major radius), e the orbital eccentricity, and $n = 2\pi/T$ the mean orbital angular velocity.

4.8 Orbital Energies

Let us now generalize our analysis to take into account the orbits of asteroids and comets about the Sun. Such orbits satisfy Equation (4.30), but can be parabolic ($e = 1$), or even hyperbolic ($e > 1$), as well as elliptical ($e < 1$). According to Equations (4.5) and (4.13), the total energy per unit mass of an object in a general orbit around the Sun is given by

$$\mathcal{E} = \frac{\dot{r}^{\,2} + r^2\,\dot{\theta}^{\,2}}{2} - \frac{G\,M}{r}. \tag{4.41}$$

It follows from Equations (4.22), (4.26), and (4.31) that

$$\mathcal{E} = \frac{h^2}{2}\left[\left(\frac{du}{d\theta}\right)^2 + u^2 - 2\,u\,u_c\right], \tag{4.42}$$

where $u = r^{-1}$, and $u_c = r_c^{-1}$. However, according to Equation (4.30),

$$u(\theta) = u_c\,(1 - e\,\cos\theta). \tag{4.43}$$

The previous two equations can be combined with Equations (4.31) and (4.34) to give

$$\mathcal{E} = \frac{u_c^2\,h^2}{2}\,(e^2-1) = \frac{G\,M}{2\,r_p}\,(e-1). \tag{4.44}$$

We conclude that *elliptical* orbits ($e < 1$) have *negative* total energies, whereas *parabolic* orbits ($e = 1$) have *zero* total energies, and *hyperbolic* orbits ($e > 1$) have *positive* total energies. This makes sense, since in a conservative system in which the potential energy at infinity is set to zero [see Equation (4.4)] we expect *bounded* orbits to have *negative* total energies, and *unbounded* orbits to have *positive* total energies—see Section 2.7. Thus, elliptical orbits, which are clearly bounded, should indeed have negative total energies,

whereas hyperbolic orbits, which are clearly unbounded, should indeed have positive total energies. Parabolic orbits are marginally bounded (*i.e.*, an object executing a parabolic orbit only just escapes from the Sun's gravitational field), and thus have zero total energy. For the special case of an elliptical orbit, whose major radius a is *finite*, we can write

$$\mathcal{E} = -\frac{G\,M}{2\,a}. \tag{4.45}$$

It follows that the energy of such an orbit is completely determined by the orbital *major radius*.

Consider an artificial satellite in an elliptical orbit around the Sun (the same considerations also apply to satellites in orbit around the Earth). At perihelion, $\dot{r} = 0$, and Equations (4.41) and (4.44) can be combined to give

$$\frac{v_t}{v_c} = \sqrt{1+e}. \tag{4.46}$$

Here, $v_t = r\,\dot{\theta}$ is the satellite's tangential velocity, and $v_c = \sqrt{G\,M/r_p}$ is the tangential velocity that it would need in order to maintain a circular orbit at the perihelion distance. Likewise, at aphelion,

$$\frac{v_t}{v_c} = \sqrt{1-e}, \tag{4.47}$$

where $v_c = \sqrt{G\,M/r_a}$ is now the tangential velocity that the satellite would need in order to maintain a circular orbit at the aphelion distance.

Suppose that our satellite is initially in a circular orbit of radius r_1, and that we wish to transfer it into a circular orbit of radius r_2, where $r_2 > r_1$. We can achieve this by temporarily placing the satellite in an elliptical orbit whose perihelion distance is r_1, and whose aphelion distance is r_2. It follows, from Equation (4.37), that the required eccentricity of the elliptical orbit is

$$e = \frac{r_2 - r_1}{r_2 + r_1}. \tag{4.48}$$

According to Equation (4.46), we can transfer our satellite from its initial circular orbit into the temporary elliptical orbit by increasing its tangential velocity (by briefly switching on the satellite's rocket motor) by a factor

$$\alpha_1 = \sqrt{1+e}. \tag{4.49}$$

We must next allow the satellite to execute half an orbit, so that it attains its aphelion distance, and then boost the tangential velocity by a factor [see Equation (4.47)]

$$\alpha_2 = \frac{1}{\sqrt{1-e}}. \tag{4.50}$$

The satellite will now be in a circular orbit at the aphelion distance, r_2. This process is illustrated in Figure 4.3. Obviously, we can transfer our satellite from a larger to a

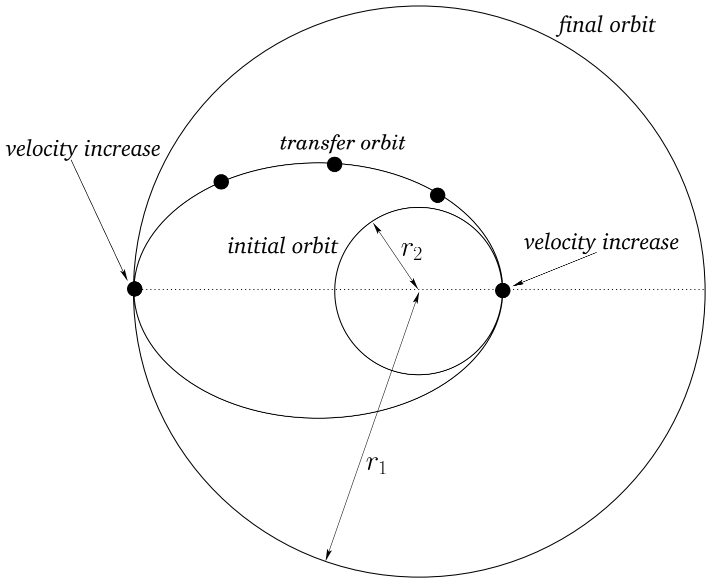


Figure 4.3: *A transfer orbit between two circular orbits.*

smaller circular orbit by performing the above process in reverse. Note, finally, from Equation (4.46), that if we increase the tangential velocity of a satellite in a circular orbit about the Sun by a factor greater than $\sqrt{2}$ then we will transfer it into a hyperbolic orbit ($e > 1$), and it will eventually escape from the Sun's gravitational field.

4.9 Kepler Problem

In a nutshell, the so-called *Kepler problem* consists of determining the radial and angular coordinates, r and θ, respectively, of an object in an elliptical, parabolic, or hyperbolic orbit about the Sun as a function of time.

Consider an object in a general Keplerian orbit about the Sun which passes through its perihelion point, $r = r_p$ and $\theta = 0$, at $t = \tau$. The constant τ is termed the *time of perihelion passage*. It follows from the previous analysis that

$$r = \frac{r_p\,(1+e)}{1+e\,\cos\theta}, \tag{4.51}$$

and

$$\mathcal{E} = \frac{\dot{r}^2}{2} + \frac{h^2}{2\,r^2} - \frac{G\,M}{r}, \tag{4.52}$$

where e, $h = \sqrt{G\,M\,r_p\,(1+e)}$, and $\mathcal{E} = G\,M\,(e-1)/(2\,r_p)$ are the orbital eccentricity, angular momentum per unit mass, and energy per unit mass, respectively. The above equation can be rearranged to give

$$\dot{r}^2 = (e-1)\,\frac{G\,M}{r_p} - (e+1)\,\frac{r_p\,G\,M}{r^2} + \frac{2\,G\,M}{r}. \tag{4.53}$$

Taking the square-root, and integrating, we obtain

$$\int_{r_p}^{r} \frac{r\,dr}{[2\,r + (e-1)\,r^2/r_p - (e+1)\,r_p]^{1/2}} = \sqrt{G\,M}\,\,(t-\tau). \tag{4.54}$$

Consider an elliptical orbit characterized by $0 < e < 1$. Let us write

$$r = \frac{r_p}{1-e}\,(1 - e\,\cos E), \tag{4.55}$$

where E is termed the *elliptic anomaly*. In fact, E is an angle which varies between $-\pi$ and $+\pi$. Moreover, the perihelion point corresponds to $E = 0$, and the aphelion point to $E = \pi$. Now,

$$dr = \frac{r_p}{1-e}\,e\,\sin E\,dE, \tag{4.56}$$

whereas

$$\begin{aligned} 2\,r + (e-1)\,\frac{r^2}{r_p} - (e+1)\,r_p &= \frac{r_p}{1-e}\,e^2\,(1-\cos^2 E) \\ &= \frac{r_p}{1-e}\,e^2\,\sin^2 E. \end{aligned} \tag{4.57}$$

Thus, Equation (4.54) reduces to

$$\int_0^E (1 - e\,\cos E)\,dE = \left(\frac{G\,M}{a^3}\right)^{1/2}(t-\tau), \tag{4.58}$$

where $a = r_p/(1-e)$. This equation can immediately be integrated to give

$$E - e\,\sin E = \mathcal{M}. \tag{4.59}$$

Here,

$$\mathcal{M} = n\,(t-\tau) \tag{4.60}$$

is termed the *mean anomaly*, $n = 2\pi/T$ is the mean orbital angular velocity, and $T = 2\pi\,(a^3/GM)^{1/2}$ is the orbital period. The mean anomaly increases *uniformly* in time at the rate of 2π radians every orbital period. Moreover, the perihelion point corresponds to $\mathcal{M} = 0$, and the aphelion point to $\mathcal{M} = \pi$. Incidentally, the angle θ, which determines the true angular location of the orbiting object, is sometimes termed the *true anomaly*. Equation (4.59), which is known as *Kepler's equation*, is a transcendental equation that does not possess a simple analytic solution. Fortunately, it is fairly straightforward to solve numerically. For instance, using an iterative approach, if E_n is the nth guess then

$$E_{n+1} = \mathcal{M} + e\,\sin E_n. \tag{4.61}$$

The above iteration scheme converges very rapidly (except in the limit $e \to 1$).

Equations (4.51) and (4.55) can be combined to give

$$\cos\theta = \frac{\cos E - e}{1 - e\,\cos E}. \tag{4.62}$$

Thus,

$$1 + \cos\theta = 2\,\cos^2(\theta/2) = \frac{2\,(1-e)\,\cos^2(E/2)}{1 - e\,\cos E}, \tag{4.63}$$

and

$$1 - \cos\theta = 2\,\sin^2(\theta/2) = \frac{2\,(1+e)\,\sin^2(E/2)}{1 - e\,\cos E}. \tag{4.64}$$

The previous two equations imply that

$$\tan(\theta/2) = \left(\frac{1+e}{1-e}\right)^{1/2}\tan(E/2). \tag{4.65}$$

We conclude that, in the case of an elliptical orbit, the solution of the Kepler problem reduces to the solution of the following three equations:

$$E - e\,\sin E = \mathcal{M}, \tag{4.66}$$

$$r = a\,(1 - e\,\cos E), \tag{4.67}$$

$$\tan(\theta/2) = \left(\frac{1+e}{1-e}\right)^{1/2}\tan(E/2). \tag{4.68}$$

Here, $T = 2\pi\,(a^3/G\,M)^{1/2}$ and $a = r_p/(1-e)$. Incidentally, it is clear that if $t \to t + T$ then $\mathcal{M} \to \mathcal{M} + 2\pi$, $E \to E + 2\pi$, and $\theta \to \theta + 2\pi$. In other words, the motion is periodic with period T. Note that since the Planets all have low eccentricity orbits, characterized by $e \ll 1$, Equations (4.66)–(4.68) can be usefully solved via series expansion in e to give

$$E = M + e\,\sin\mathcal{M} + \frac{e^2}{2}\,\sin 2\mathcal{M} + \frac{e^3}{8}\,(3\,\sin 3\mathcal{M} - \sin\mathcal{M}) + \mathcal{O}(e^4), \tag{4.69}$$

$$\theta = \mathcal{M} + 2\,e\,\sin\mathcal{M} + \frac{5\,e^2}{4}\,\sin 2\mathcal{M} + \frac{e^3}{12}\,(13\,\sin 3\mathcal{M} - 3\,\sin\mathcal{M}) + \mathcal{O}(e^4), \tag{4.70}$$

$$\frac{r}{a} = 1 - e\,\cos\mathcal{M} + \frac{e^2}{2}\,(1 - \cos 2\mathcal{M}) + \frac{3\,e^3}{8}\,(\cos\mathcal{M} - \cos 3\mathcal{M}) + \mathcal{O}(e^4). \tag{4.71}$$

For the case of a parabolic orbit, characterized by $e = 1$, similar analysis to the above yields:

$$P + \frac{P^3}{3} = \left(\frac{G\,M}{2\,r_p^3}\right)^{1/2}(t - \tau), \tag{4.72}$$

$$r = r_p\,(1 + P^2), \tag{4.73}$$

$$\tan(\theta/2) = P. \tag{4.74}$$

Here, P is termed the *parabolic anomaly,* and varies between $-\infty$ and $+\infty$, with the perihelion point corresponding to $P = 0$. Note that Equation (4.72) is a cubic equation, possessing a single real root, which can, in principle, be solved analytically. However, a numerical solution is generally more convenient.

Finally, for the case of a hyperbolic orbit, characterized by $e > 1$, we obtain:

$$e\,\sinh H - H = \left(\frac{G\,M}{a^3}\right)^{1/2}(t-\tau), \tag{4.75}$$

$$r = a\,(e\,\cosh H - 1), \tag{4.76}$$

$$\tan(\theta/2) = \left(\frac{e+1}{e-1}\right)^{1/2}\tanh(H/2). \tag{4.77}$$

Here, H is termed the *hyperbolic anomaly,* and varies between $-\infty$ and $+\infty$, with the perihelion point corresponding to $H = 0$. Moreover, $a = r_p/(e-1)$. As in the elliptical case, Equation (4.75) is a transcendental equation which is most easily solved numerically.

4.10 Orbital Elements

The previous analysis suffices when considering a single planet orbiting around the Sun. However, it becomes inadequate when dealing with multiple planets whose orbital planes and perihelion directions do not necessarily coincide. Incidentally, for the time being, we are neglecting interplanetary gravitational interactions, which allows us to assume that each planet executes an *independent* Keplerian orbit about the Sun.

Let us characterize all planetary orbits using a common Cartesian coordinate system (X, Y, Z), centered on the Sun—see Fig. 4.4. The X-Y plane defines a *reference plane,* which is chosen to be the *ecliptic plane* (*i.e.*, the plane of the Earth's orbit), with the Z-axis pointing towards the northern ecliptic pole (*i.e.*, the direction normal to the ecliptic plane in a northward sense). Likewise, the X-axis defines a *reference direction,* which is chosen to point in the direction of the vernal equinox at the same epoch (*i.e.*, the point in the sky at which the ecliptic plane appears to pass through the extension of the Earth's equatorial plane from south to north). Suppose that the plane of a given planetary orbit is inclined at an angle I to the reference plane. The point at which this orbit crosses the reference plane in the direction of increasing Z is termed its *ascending node*. The angle Ω subtended between the reference direction and the direction of the ascending node is termed the *longitude of the ascending node*. Finally, the angle, ω, subtended between the direction of the ascending node and the direction of the orbit's perihelion is termed the *argument of the perihelion*.

Let us define a second Cartesian coordinate system (x, y, z), also centered on the Sun. Let the x-y plane coincide with the plane of a particular orbit, and let the x-axis point towards the orbit's perihelion point. Clearly, we can transform from the (x, y, z) system to the (X, Y, Z) system via a series of three rotations of the coordinate system. First, a rotation

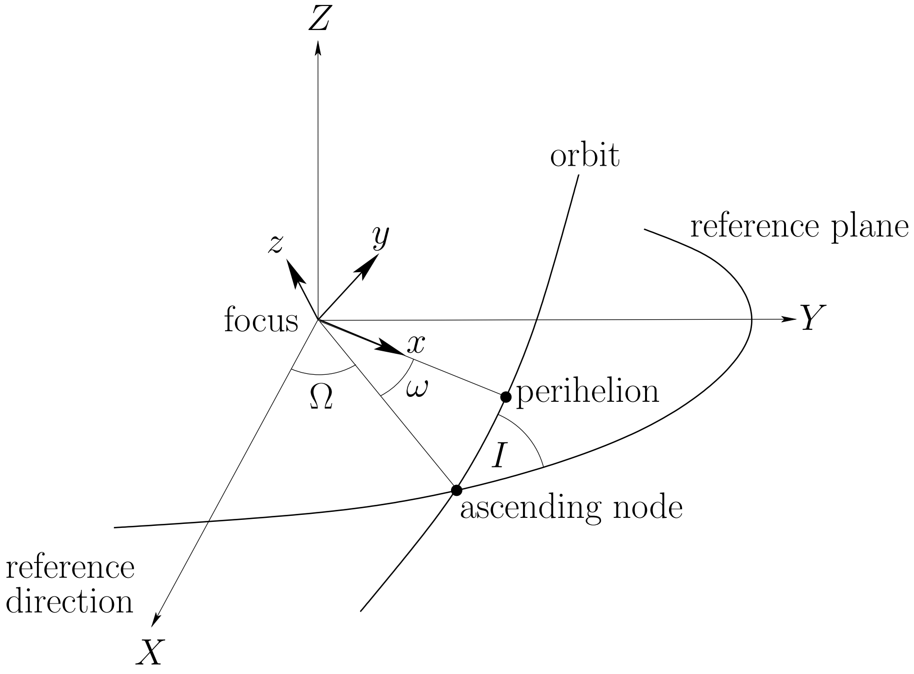


Figure 4.4: *A general planetary orbit.*

through an angle ω about the z-axis (looking down the axis). Second, a rotation through an angle I about the new x-axis. Finally, a rotation through an angle Ω about the new z-axis. It, thus, follows from standard coordinate transformation theory (see Section A.5) that

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos I & -\sin I \\ 0 & \sin I & \cos I \end{pmatrix} \begin{pmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.78}$$

However, $x = r\,\cos\theta$, $y = r\,\sin\theta$, and $z = 0$. Hence,

$$X = r\,[\cos\Omega\,\cos(\omega+\theta) - \sin\Omega\,\sin(\omega+\theta)\,\cos I], \tag{4.79}$$

$$Y = r\,[\sin\Omega\,\cos(\omega+\theta) + \cos\Omega\,\sin(\omega+\theta)\,\cos I], \tag{4.80}$$

$$Z = r\,\sin(\omega+\theta)\,\sin I. \tag{4.81}$$

Thus, a general planetary orbit is determined by Equations (4.66)–(4.68) and (4.79)–(4.81), and is therefore parameterized by *six* orbital elements: the minor radius, a; the eccentricity, e; the time of perihelion passage, τ; the inclination, I; the longitude of the ascending node, Ω; and the argument of the perihelion, ω. (The orbital angular velocity, in radians per year, is $n = 2\pi/a^{3/2}$, where a is measured in astronomical units—the major radius of the Earth's orbit is defined to be 1 AU.) The argument of the perihelion is

Planet	$a(\text{au})$	e	$\bar{\lambda}_0(^\circ)$	$I(^\circ)$	$\Omega(^\circ)$	$\varpi(^\circ)$	m/M
Mercury	0.3871	0.20564	252.25	7.006	48.34	77.46	1.659×10^{-7}
Venus	0.7233	0.00676	181.98	3.398	76.67	131.77	2.447×10^{-6}
Earth	1.0000	0.01673	100.47	0.000	344.89	102.93	3.039×10^{-6}
Mars	1.5237	0.09337	355.43	1.852	49.71	336.08	3.226×10^{-7}
Jupiter	5.2025	0.04854	34.33	1.299	100.29	14.21	9.542×10^{-4}
Saturn	9.5415	0.05551	50.08	2.494	113.64	93.00	2.857×10^{-4}
Uranus	19.188	0.04686	314.20	0.773	73.96	173.41	4.353×10^{-5}
Neptune	30.070	0.00895	304.22	1.770	131.79	46.00	5.165×10^{-5}

Table 4.1: *Orbital elements of the major planets at J2000. Here,* a *is the major radius (in astronomical units),* e *the eccentricity,* $\bar{\lambda}_0$ *the mean longitude at epoch,* I *the inclination,* Ω *the longitude of the ascending node, and* ϖ *the longitude of the perihelion. The last column gives the mass of each planet as a fraction of the solar mass. [E.M. Standish,* et al.*, Orbital Ephemeris of the Sun, Moon, and Planets, in* Explanatory Supplement to the Astronomical Almanac*, ed. P.K. Seidelmann (University Science Books, Mill Valley CA, 1992.)]*

sometimes replaced by

$$\varpi = \Omega + \omega, \tag{4.82}$$

which is termed the *longitude of the perihelion*. Likewise, the time of perihelion passage, τ, is often replaced by the mean longitude at $t = 0$, where the *mean longitude* is defined

$$\bar{\lambda} = \varpi + M = \varpi + n\,(t - \tau). \tag{4.83}$$

Thus, if $\bar{\lambda}_0$ denotes the mean longitude at epoch ($t = 0$), then

$$\bar{\lambda} = \bar{\lambda}_0 + n\,t, \tag{4.84}$$

where $\bar{\lambda}_0 = \varpi - n\,\tau$. The orbital elements of the major planets at the epoch J2000 (*i.e.*, at 00:00 UT on January 1st, 2000) are given in Table 4.1.

The *heliocentric* (*i.e.*, as seen from the Sun) position of a planet is most conveniently expressed in terms of its ecliptic longitude, λ, and ecliptic latitude, β. This type of longitude and latitude is referred to the ecliptic plane, with the Sun as the origin. It follows that

$$\tan\lambda = \frac{Y}{X}, \tag{4.85}$$

$$\sin\beta = \frac{Z}{\sqrt{X^2+Y^2}}, \tag{4.86}$$

where (X, Y, Z) are the heliocentric Cartesian coordinates of the planet.

4.11 Binary Star Systems

Approximately half of the stars in our galaxy are members of so-called *binary star systems*. Such systems consist of two stars, of mass m_1 and m_2, and position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively, orbiting about their common center of mass. The distance separating the stars is always much less than the distance to the nearest neighbor star. Hence, a binary star system can be treated as a two-body dynamical system to a very good approximation.

In a binary star system, the gravitational force which the first star exerts on the second is

$$\mathbf{f} = -\frac{G\, m_1\, m_2}{r^3}\,\mathbf{r}, \tag{4.87}$$

where $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. As we have seen, a two-body system can be reduced to an equivalent one-body system whose equation of motion is of the form (2.76), where $\mu = m_1\, m_2/(m_1 + m_2)$. Hence, in this particular case, we can write

$$\frac{m_1\, m_2}{m_1 + m_2}\,\frac{d^2\mathbf{r}}{dt^2} = -\frac{G\, m_1\, m_2}{r^3}\,\mathbf{r}, \tag{4.88}$$

which gives

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{G\, M}{r^3}\,\mathbf{r}, \tag{4.89}$$

where

$$M = m_1 + m_2. \tag{4.90}$$

Equation (4.89) is identical to Equation (4.2), which we have already solved. Hence, we can immediately write down the solution:

$$\mathbf{r} = (r\,\cos\theta,\; r\,\sin\theta,\; 0), \tag{4.91}$$

where

$$r = \frac{a\,(1-e^2)}{1 - e\,\cos\theta}, \tag{4.92}$$

and

$$\frac{d\theta}{dt} = \frac{h}{r^2}, \tag{4.93}$$

with

$$a = \frac{h^2}{(1-e^2)\,G\,M}. \tag{4.94}$$

Here, h is a constant, and we have aligned our Cartesian axes so that the plane of the orbit coincides with the x-y plane. According to the above solution, the second star executes a Keplerian elliptical orbit, with major radius a and eccentricity e, relative to the first star, and *vice versa*. From Equation (4.33), the period of revolution, T, is given by

$$T = \sqrt{\frac{4\pi^2\, a^3}{G\, M}}. \tag{4.95}$$

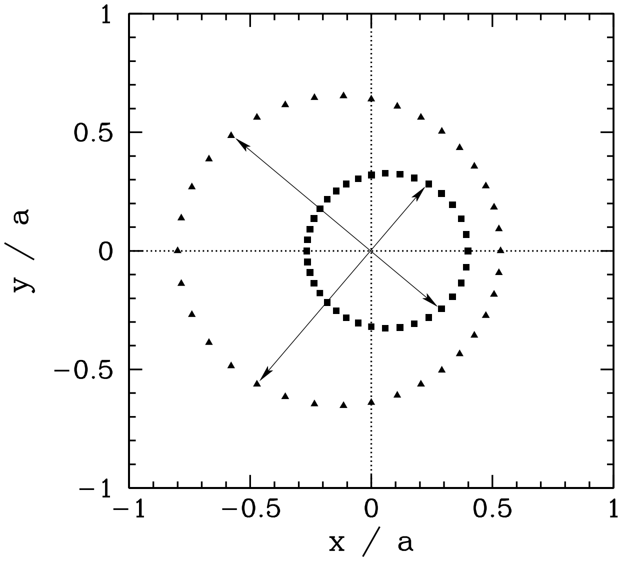


Figure 4.5: *An example binary star orbit.*

Moreover, if $n = 2\pi/T$ then

$$n = \frac{\sqrt{G\,M}}{a^{3/2}}. \tag{4.96}$$

In the *inertial* frame of reference whose origin always coincides with the center of mass—the so-called *center of mass frame*—the position vectors of the two stars are

$$\mathbf{r}_1 = -\frac{m_2}{m_1 + m_2}\,\mathbf{r}, \tag{4.97}$$

$$\mathbf{r}_2 = \frac{m_1}{m_1 + m_2}\,\mathbf{r}, \tag{4.98}$$

where $\mathbf{r}$ is specified above. Figure 4.5 shows an example binary star orbit, in the center of mass frame, calculated with $m_1/m_2 = 0.5$ and $e = 0.2$. Here, the triangles and squares denote the positions of the first and second star, respectively (which are always diagrammatically opposite one another, relative to the origin, as indicated by the arrows). It can be seen that both stars execute elliptical orbits about their common center of mass.

Binary star systems have been very useful to astronomers, since it is possible to determine the masses of both stars in such a system by careful observation. The *sum* of the masses of the two stars, $M = m_1 + m_2$, can be found from Equation (4.95) after a measurement of the major radius, a (which is the mean of the greatest and smallest distance apart of the two stars during their orbit), and the orbital period, T. The *ratio* of the masses of the two stars, m_1/m_2, can be determined from Equations (4.97) and (4.98) by observing the fixed ratio of the relative distances of the two stars from the common center of mass about which they both appear to rotate. Obviously, given the sum of the masses, and the ratio of the masses, the individual masses themselves can then be calculated.

4.12 Exercises

4.1. Halley's comet has an orbital eccentricity of $e = 0.967$ and a perihelion distance of $55,000,000$ miles. Find the orbital period, and the comet's speed at perihelion and aphelion.

4.2. A comet is first seen at a distance of d astronomical units (1 astronomical unit is the mean Earth-Sun distance) from the Sun, and is traveling with a speed of q times the Earth's mean speed. Show that the orbit of the comet is hyperbolic, parabolic, or elliptical, depending on whether the quantity $q^2\,d$ is greater than, equal to, or less than 2, respectively.

4.3. Consider a planet in a Keplerian orbit of major radius a and eccentricity e about the Sun. Suppose that the eccentricity of the orbit is small (*i.e.*, $0 < e \ll 1$), as is indeed the case for all of the planets except Mercury and Pluto. Demonstrate that, to first-order in e, the orbit can be approximated as a *circle* whose center is shifted a distance $e\,a$ from the Sun, and that the planet's angular motion appears *uniform* when viewed from a point (called the Equant) which is shifted a distance $2\,e\,a$ from the Sun, in the same direction as the center of the circle. This theorem is the basis of the Ptolomaic model of planetary motion.

4.4. How long (in days) does it take the Sun-Earth radius vector to rotate through 90°, starting at the perihelion point? How long does it take starting at the aphelion point? The period and eccentricity of the Earth's orbit are $T = 365.24$ days, and $e = 0.01673$, respectively.

4.5. Solve the Kepler problem for a parabolic orbit to obtain Equations (4.72)–(4.74).

4.6. Solve the Kepler problem for a hyperbolic orbit to obtain Equations (4.75)–(4.77).

4.7. A comet is in a parabolic orbit lying in the plane of the Earth's orbit. Regarding the Earth's orbit as a circle of radius a, show that the points where the comet intersects the Earth's orbit are given by

$$\cos\theta = -1 + \frac{2\,p}{a},$$

where p is the perihelion distance of the comet, defined at $\theta = 0$. Show that the time interval that the comet remains inside the Earth's orbit is the faction

$$\frac{2^{1/2}}{3\pi}\left(\frac{2\,p}{a} + 1\right)\left(1 - \frac{p}{a}\right)^{1/2}$$

of a year, and that the maximum value of this time interval is $2/3\pi$ year, or about 11 weeks.

5 Orbits in Central Force-Fields

5.1 Introduction

This chapter examines the motion of a celestial body in a general central potential (*i.e.*, a potential that is a function of r, but does not necessarily vary as $1/r$).

5.2 Motion in a General Central Force-Field

Consider the motion of an object in a general (attractive) central force-field characterized by the potential energy *per unit mass* function $V(r)$. Since the force-field is central, it still remains true that

$$h = r^2\,\dot{\theta} \tag{5.1}$$

is a constant of the motion (see Section 4.5). As is easily demonstrated, Equation (4.28) generalizes to

$$\frac{d^2u}{d\theta^2} + u = -\frac{1}{h^2}\frac{dV}{du}, \tag{5.2}$$

where $u = r^{-1}$.

Suppose, for instance, that we wish to find the potential $V(r)$ that causes an object to execute the spiral orbit

$$r = r_0\,\theta^{\,2}. \tag{5.3}$$

Substitution of $u = (r_0\,\theta^2)^{-1}$ into Equation (5.2) yields

$$\frac{dV}{du} = -h^2\left(6\,r_0\,u^2 + u\right). \tag{5.4}$$

Integrating, we obtain

$$V(u) = -h^2\left(2\,r_0\,u^3 + \frac{u^2}{2}\right), \tag{5.5}$$

or

$$V(r) = -h^2\left(\frac{2\,r_0}{r^3} + \frac{1}{2\,r^2}\right). \tag{5.6}$$

In other words, the spiral pattern (5.3) is obtained from a mixture of an inverse-square and inverse-cube potential.

5.3 Motion in a Nearly Circular Orbit

In principle, a circular orbit is a possible orbit for any attractive central force. However, not all such forces result in *stable* circular orbits. Let us now consider the stability of circular

orbits in a general central force-field. Equation (4.25) generalizes to

$$\ddot{r} - \frac{h^2}{r^3} = f(r), \tag{5.7}$$

where $f(r)$ is the radial force *per unit mass*. For a circular orbit, $\ddot{r} = 0$, and the above equation reduces to

$$-\frac{h^2}{r_c^3} = f(r_c), \tag{5.8}$$

where r_c is the radius of the orbit.

Let us now consider *small* departures from circularity. Let

$$x = r - r_c. \tag{5.9}$$

Equation (5.7) can be written

$$\ddot{x} - \frac{h^2}{(r_c + x)^3} = f(r_c + x). \tag{5.10}$$

Expanding the two terms involving $r_c + x$ as power series in x/r_c, and keeping all terms up to first order, we obtain

$$\ddot{x} - \frac{h^2}{r_c^3}\left(1 - 3\,\frac{x}{r_c}\right) = f(r_c) + f'(r_c)\,x, \tag{5.11}$$

where $'$ denotes a derivative. Making use of Equation (5.8), the above equation reduces to

$$\ddot{x} + \left[-\frac{3\,f(r_c)}{r_c} - f'(r_c)\right] x = 0. \tag{5.12}$$

If the term in square brackets is *positive* then we obtain a simple harmonic equation, which we already know has bounded solutions (see Section 2.8)—*i.e.*, the orbit is *stable* to small perturbations. On the other hand, if the term is square brackets is *negative* then we obtain an equation whose solutions grow exponentially in time (see Section 2.8)—*i.e.*, the orbit is *unstable* to small oscillations. Thus, the stability criterion for a circular orbit of radius r_c in a central force-field characterized by a radial force (per unit mass) function $f(r)$ is

$$f(r_c) + \frac{r_c}{3}\,f'(r_c) < 0. \tag{5.13}$$

For example, consider an attractive power-law force function of the form

$$f(c) = -c\,r^n, \tag{5.14}$$

where $c > 0$. Substituting into the above stability criterion, we obtain

$$-c\,r_c^n - \frac{c\,n}{3}\,r_c^n < 0, \tag{5.15}$$

or

$$n > -3. \tag{5.16}$$

We conclude that circular orbits in attractive central force-fields that decay faster than r^{-3} are unstable. The case $n = -3$ is special, since the first-order terms in the expansion of Equation (5.10) cancel out exactly, and it is necessary to retain the second-order terms. Doing this, it is easily demonstrated that circular orbits are also unstable for inverse-cube ($n = -3$) forces.

An *apsis* (plural, apsides) is a point in an orbit at which the radial distance, r, assumes either a *maximum* or a *minimum* value. Thus, the perihelion and aphelion points are the apsides of planetary orbits. The angle through which the radius vector rotates in going between two consecutive apsides is called the *apsidal angle*. Thus, the apsidal angle for elliptical orbits in an inverse-square force-field is π.

For the case of stable nearly circular orbits, we have seen that r oscillates sinusoidally about its mean value, r_c. Indeed, it is clear from Equation (5.12) that the period of the oscillation is

$$T = \frac{2\pi}{[-3\,f(r_c)/r_c - f'(r_c)]^{1/2}}. \tag{5.17}$$

The apsidal angle is the amount by which θ increases in going between a maximum and a minimum of r. The time taken to achieve this is clearly $T/2$. Now, $\dot{\theta} = h/r^2$, where h is a constant of the motion, and r is almost constant. Thus, $\dot{\theta}$ is approximately constant. In fact,

$$\dot{\theta} \simeq \frac{h}{r_c^2} = \left[-\frac{f(r_c)}{r_c}\right]^{1/2}, \tag{5.18}$$

where use has been made of Equation (5.8). Thus, the apsidal angle, ψ, is given by

$$\psi = \frac{T}{2}\,\dot{\theta} = \pi\left[3 + r_c\,\frac{f'(r_c)}{f(r_c)}\right]^{-1/2} \tag{5.19}$$

For the case of attractive power-law central forces of the form $f(r) = -c\,r^n$, where $c > 0$, the apsidal angle becomes

$$\psi = \frac{\pi}{(3+n)^{1/2}}. \tag{5.20}$$

Now, it should be clear that if an orbit is going to close on itself then the apsidal angle needs to be a *rational* fraction of 2π. There are, in fact, only two small integer values of the power-law index, n, for which this is the case. As we have seen, for an inverse-square force law (*i.e.*, $n = -2$), the apsidal angle is π. For a linear force law (*i.e.*, $n = 1$), the apsidal angle is $\pi/2$. However, for quadratic (*i.e.*, $n = 2$) or cubic (*i.e.*, $n = 3$) force laws, the apsidal angle is an *irrational* fraction of 2π, which means that non-circular orbits in such force-fields never close on themselves.

5.4 Perihelion Precession of the Planets

The Solar System consists of eight major planets (Mercury to Neptune) moving around the Sun in slightly elliptical orbits that are approximately co-planar with one another. According to Chapter 4, if we neglect the relatively weak interplanetary gravitational interactions then the perihelia of the various planets (*i.e.*, the points on their orbits at which they are closest to the Sun) remain *fixed* in space. However, once these interactions are taken into account, it turns out that the planetary perihelia all slowly *precess* around the Sun. We can calculate the approximate rate of perihelion precession of a given planet by treating the other planets as *uniform concentric rings*, centered on the Sun, of mass equal to the planetary mass, and radius equal to the mean orbital radius. This method of calculation, which is due to Gauss, is equivalent to averaging the interplanetary gravitational interactions over the orbits of the other planets. It is reasonable to do this because the precession period in question is very much longer than the orbital period of any planet in the Solar System. Thus, by treating the other planets as rings, we can calculate the mean gravitational perturbation due to these planets, and, thereby, determine the desired precession rate.

We can conveniently index the planets in the Solar System such that Mercury is planet 1, and Neptune planet 8. Let the M_i and the R_i, for $i = 1, 8$, be the planetary masses and orbital radii, respectively. Furthermore, let M_0 be the mass of the Sun. It follows, from Section 3.7, that the gravitational potential generated at the ith planet by the Sun and the other planets is

$$\begin{aligned}\Phi(R_i) &= -\frac{G\,M_0}{R_i} - G\sum_{j<i}\frac{M_j}{R_i}\left[1+\frac{1}{4}\left(\frac{R_j}{R_i}\right)^2+\frac{9}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]\\ &\quad -G\sum_{j>i}\frac{M_j}{R_j}\left[1+\frac{1}{4}\left(\frac{R_i}{R_j}\right)^2+\frac{9}{64}\left(\frac{R_i}{R_j}\right)^4+\cdots\right].\end{aligned} \tag{5.21}$$

Now, the radial force per unit mass acting on the ith planet is written $f(R_i) = -d\Phi(R_i)/dr$, giving

$$\begin{aligned}f(R_i) &= -\frac{G\,M_0}{R_i^2} - \frac{G}{R_i^2}\sum_{j<i}M_j\left[1+\frac{3}{4}\left(\frac{R_j}{R_i}\right)^2+\frac{45}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]\\ &\quad +\frac{G}{R_i^2}\sum_{j>i}M_j\left(\frac{R_i}{R_j}\right)\left[\frac{1}{2}\left(\frac{R_i}{R_j}\right)^2+\frac{9}{16}\left(\frac{R_i}{R_j}\right)^4+\cdots\right].\end{aligned} \tag{5.22}$$

Hence, we obtain

$$R_i\,f'(R_i) = \frac{2\,G\,M_0}{R_i^2} + \frac{G}{R_i^2}\sum_{j<i}M_j\left[2+3\left(\frac{R_j}{R_i}\right)^2+\frac{135}{32}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]$$

$$+\frac{G}{R_i^2}\sum_{j>i} M_j\left(\frac{R_i}{R_j}\right)\left[\frac{1}{2}\left(\frac{R_i}{R_j}\right)^2+\frac{27}{16}\left(\frac{R_i}{R_j}\right)^4+\cdots\right], \tag{5.23}$$

where $' \equiv d/dr$. It follows that

$$\left[3+\frac{R_i\,f'(R_i)}{f(R_i)}\right]^{-1/2} = 1+\frac{3}{4}\sum_{j<i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_j}{R_i}\right)^2\left[1+\frac{15}{8}\left(\frac{R_j}{R_i}\right)^2+\frac{175}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]$$
$$+\frac{3}{4}\sum_{j>i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_i}{R_j}\right)^3\left[1+\frac{15}{8}\left(\frac{R_i}{R_j}\right)^2+\frac{175}{64}\left(\frac{R_i}{R_j}\right)^4+\cdots\right]. \tag{5.24}$$

Thus, according to Equation (5.19), the apsidal angle for the ith planet is

$$\psi_i = \pi+\frac{3\pi}{4}\sum_{j<i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_j}{R_i}\right)^2\left[1+\frac{15}{8}\left(\frac{R_j}{R_i}\right)^2+\frac{175}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]$$
$$+\frac{3\pi}{4}\sum_{j>i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_i}{R_j}\right)^3\left[1+\frac{15}{8}\left(\frac{R_i}{R_j}\right)^2+\frac{175}{64}\left(\frac{R_i}{R_j}\right)^4+\cdots\right]. \tag{5.25}$$

Hence, the perihelion of the ith planet advances by

$$\delta\psi_i = \frac{3\pi}{2}\sum_{j<i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_j}{R_i}\right)^2\left[1+\frac{15}{8}\left(\frac{R_j}{R_i}\right)^2+\frac{175}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]$$
$$+\frac{3\pi}{2}\sum_{j>i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_i}{R_j}\right)^3\left[1+\frac{15}{8}\left(\frac{R_i}{R_j}\right)^2+\frac{175}{64}\left(\frac{R_i}{R_j}\right)^4+\cdots\right] \tag{5.26}$$

radians per revolution around the Sun. Now, the time for one revolution is $T_i = 2\pi/\omega_i$, where $\omega_i^2 = G\,M_0/R_i^3$. Thus, the rate of perihelion precession, in *arc seconds per year*, is given by

$$\delta\dot{\psi}_i = \frac{75}{T_i(\mathrm{yr})}\left\{\sum_{j<i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_j}{R_i}\right)^2\left[1+\frac{15}{8}\left(\frac{R_j}{R_i}\right)^2+\frac{175}{64}\left(\frac{R_j}{R_i}\right)^4+\cdots\right]\right.$$
$$\left.+\sum_{j>i}\left(\frac{M_j}{M_0}\right)\left(\frac{R_i}{R_j}\right)^3\left[1+\frac{15}{8}\left(\frac{R_i}{R_j}\right)^2+\frac{175}{64}\left(\frac{R_i}{R_j}\right)^4+\cdots\right]\right\}. \tag{5.27}$$

Table 5.2 and Figure 5.1 compare the observed perihelion precession rates with the theoretical rates calculated from Equation (5.27) and the planetary data given in Table 5.1. It can be seen that there is excellent agreement between the two, except for the planet Venus. The main reason for this is that Venus has an unusually low eccentricity ($e = 0.0068$), which renders its perihelion point extremely sensitive to small perturbations.

Planet	M/M_0	T(yr)	R(au)
Mercury	1.66×10^{-7}	0.241	0.387
Venus	2.45×10^{-6}	0.615	0.723
Earth	3.04×10^{-6}	1.000	1.00
Mars	3.23×10^{-7}	1.881	1.52
Jupiter	9.55×10^{-4}	11.86	5.20
Saturn	2.86×10^{-4}	29.46	9.54
Uranus	4.36×10^{-5}	84.01	19.19
Neptune	5.18×10^{-5}	164.8	30.07

Table 5.1: *Data for the major planets in the Solar System, giving the planetary mass relative to that of the Sun, the orbital period in years, and the mean orbital radius relative to that of the Earth. [E.M. Standish,* et al.*, Orbital Ephemeris of the Sun, Moon, and Planets, in* Explanatory Supplement to the Astronomical Almanac*, ed. P.K. Seidelmann (University Science Books, Mill Valley CA, 1992.)]*

Planet	$(\delta\dot{\Psi})_{obs}$	$(\delta\dot{\Psi})_{th}$
Mercury	5.75	5.50
Venus	2.04	10.75
Earth	11.45	11.87
Mars	16.28	17.60
Jupiter	6.55	7.42
Saturn	19.50	18.36
Uranus	3.34	2.72
Neptune	0.36	0.65

Table 5.2: *The observed perihelion precession rates of the planets compared with the theoretical precession rates calculated from Equation (5.27) and Table 5.1. The precession rates are in arc seconds per year. (*ibid.*)*

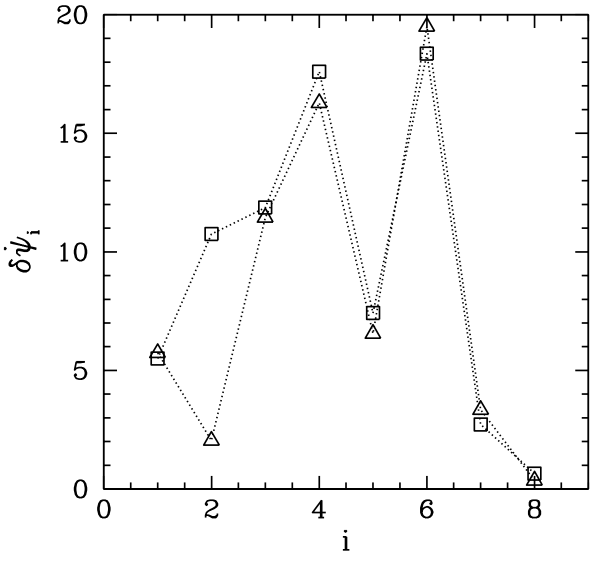


Figure 5.1: *The triangular points show the observed perihelion precession rates of the major planets in the Solar System, whereas the square points show the theoretical rates calculated from Equation (5.27) and Table 5.1. The precession rates are in arc seconds per year.*

5.5 Perihelion Precession of Mercury

If the calculation described in the previous section is carried out more accurately, taking into account the slight eccentricities of the planetary orbits, as well as their small mutual inclinations, and retaining many more terms in the expansions (3.70) and (3.72), then the perihelion precession rate of the planet Mercury is found to be 5.32 arc seconds per year.[1] However, the observed precession rate is 5.75 arc seconds per year. It turns out that the cause of this discrepancy is the general relativistic correction to Newtonian gravity.

General relativity gives rise to a small correction to the force per unit mass exerted by the Sun (mass M) on a planet in a circular orbit of radius r, and angular momentum per unit mass h. This correction is due to the curvature of space in the immediate vicinity of the Sun. In fact, the modified formula for f is

$$f \simeq -\frac{G\,M}{r^2} - \frac{3\,G\,M\,h^2}{c^2\,r^4}, \tag{5.28}$$

where c is the velocity of light in vacuum. It follows that

$$\frac{r\,f'}{f} = -2\left(1 + \frac{3\,h^2}{c^2\,r^2} + \cdots\right). \tag{5.29}$$

Hence, from Equation (5.19), the apsidal angle is

$$\psi \simeq \pi\left(1 + \frac{3\,h^2}{c^2\,r^2}\right). \tag{5.30}$$

[1]M.G. Stewart, American Jou. Physics **73**, 730 (2005).

Thus, the perihelion advances by

$$\delta\psi \simeq \frac{6\pi\,G\,M}{c^2\,r} \tag{5.31}$$

radians per revolution due to the general relativistic correction to Newtonian gravity. Here, use has been made of $h^2 = G\,M\,r$. It follows that the rate of perihelion precession due to the general relativistic correction is

$$\delta\dot{\psi} \simeq \frac{0.0383}{R\,T} \tag{5.32}$$

arc seconds per year, where R is the mean orbital radius in mean Earth orbital radii (*i.e.*, astronomical units), and T is the orbital period in years. Hence, from Table 5.1, the general relativistic contribution to $\delta\dot{\psi}$ for Mercury is 0.41 arc seconds per year. It is easily demonstrated that the corresponding contribution is negligible for the other planets in the Solar System. If the above calculation is carried out sightly more accurately, taking the eccentricity of Mercury's orbit into account, then the general relativistic contribution to $\delta\dot{\psi}$ becomes 0.43 arc seconds per year. It follows that the total perihelion precession rate for Mercury is $5.32 + 0.43 = 5.75$ arc seconds per year. This is in exact agreement with the observed precession rate. Indeed, the ability of general relativity to explain the discrepancy between the observed perihelion precession rate of Mercury, and that calculated from Newtonian mechanics, was one of the first major successes of this theory.

5.6 Exercises

5.1. Prove that in the case of a central force varying inversely as the cube of the distance

$$r^2 = A\,t^2 + B\,t + C,$$

where A, B, C are constants.

5.2. The orbit of a particle moving in a central field is a circle passing through the origin, namely $r = r_0\,\cos\theta$. Show that the force law is inverse-fifth power.

5.3. A particle moving in a central field describes a spiral orbit $r = r_0\,\exp(k\,\theta)$. Show that the force law is inverse-cube, and that θ varies logarithmically with t. Show that there are two other possible types of orbit in this force-field, and give their equations.

5.4. A particle moves in a spiral orbit given by $r = a\,\theta$. Suppose that θ increases linearly with t. Is the force acting on the particle central in nature? If not, determine how θ would have to vary with t in order to make the force central.

5.5. A particle moves in a circular orbit of radius r_0 in an attractive central force-field of the form $f(r) = -c\,\exp(-r/a)/r^2$, where $c > 0$ and $a > 0$. Demonstrate that the orbit is only stable provided that $r_0 < a$.

5.6. A particle moves in a circular orbit in an attractive central force-field of the form $f(r) = -c\,r^{-3}$, where $c > 0$. Demonstrate that the orbit is unstable to small perturbations.

5.7. If the Solar System were embedded in a uniform dust cloud, what would the apsidal angle of a planet be for motion in a nearly circular orbit? Express your answer in terms of the ratio of the mass of dust contained in a sphere, centered on the Sun, whose radius is that of the orbit, to the mass of the Sun. This model was once suggested as a possible explanation for the advance of the perihelion of Mercury.

5.8. The potential energy per unit mass of a particle in the gravitational field of an oblate spheroid, like the Earth, is

$$V(r) = -\frac{G\,M}{r}\left(1 + \frac{\epsilon}{r^2}\right),$$

where r refers to distances in the equatorial plane, M is the Earth's mass, and $\epsilon = (2/5)\,R\,\Delta R$. Here, $R = 4000$ mi is the Earth's equatorial radius, and $\Delta R = 13$ mi the difference between the equatorial and polar radii. Find the apsidal angle for a satellite moving in a nearly circular orbit in the equatorial plane of the Earth.

6 Rotating Reference Frames

6.1 Introduction

As we saw in Chapter 2, Newton's second law of motion is only valid in *inertial* frames of reference. However, it is sometimes convenient to observe motion in non-inertial *rotating* reference frames. For instance, it is most convenient for us to observe the motions of objects close to the Earth's surface in a reference frame which is *fixed* relative to this surface. Such a frame is non-inertial in nature, since it *accelerates* with respect to a standard inertial frame due to the Earth's daily rotation about its axis. (Note that the accelerations of this frame due to the Earth's orbital motion about the Sun, or the Sun's orbital motion about the galactic center, *etc*., are negligible compared to the acceleration due to the Earth's axial rotation.) Let us investigate motion in such rotating reference frames.

6.2 Rotating Reference Frames

Suppose that a given object has position vector $\mathbf{r}$ in some *inertial* (*i.e.*, non-rotating) reference frame. Let us observe the motion of this object in a *non-inertial* reference frame which *rotates* with *constant* angular velocity $\boldsymbol{\Omega}$ about an axis passing through the origin of the inertial frame. Suppose, first of all, that our object appears *stationary* in the rotating reference frame. Hence, in the non-rotating frame, the object's position vector $\mathbf{r}$ will appear to *precess* about the origin with angular velocity $\boldsymbol{\Omega}$. It follows, from Equation (A.49), that in the non-rotating reference frame

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\Omega} \times \mathbf{r}. \tag{6.1}$$

Suppose, now, that our object appears to move in the rotating reference frame with instantaneous velocity $\mathbf{v}'$. It is fairly obvious that the appropriate generalization of the above equation is simply

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}' + \boldsymbol{\Omega} \times \mathbf{r}. \tag{6.2}$$

Let and d/dt and d/dt' denote apparent time derivatives in the non-rotating and rotating frames of reference, respectively. Since an object which is stationary in the rotating reference frame appears to move in the non-rotating frame, it is clear that $d/dt \neq d/dt'$. Writing the apparent velocity, $\mathbf{v}'$, of our object in the rotating reference frame as $d\mathbf{r}/dt'$, the above equation takes the form

$$\frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{dt'} + \boldsymbol{\Omega} \times \mathbf{r}, \tag{6.3}$$

or

$$\frac{d}{dt} = \frac{d}{dt'} + \boldsymbol{\Omega} \times, \tag{6.4}$$

since $\mathbf{r}$ is a general position vector. Equation (6.4) expresses the relationship between apparent time derivatives in the non-rotating and rotating reference frames.

Operating on the general position vector $\mathbf{r}$ with the time derivative (6.4), we get

$$\mathbf{v} = \mathbf{v}' + \boldsymbol{\Omega} \times \mathbf{r}. \tag{6.5}$$

This equation relates the apparent velocity, $\mathbf{v} = d\mathbf{r}/dt$, of an object with position vector $\mathbf{r}$ in the non-rotating reference frame to its apparent velocity, $\mathbf{v}' = d\mathbf{r}/dt'$, in the rotating reference frame.

Operating twice on the position vector $\mathbf{r}$ with the time derivative (6.4), we obtain

$$\mathbf{a} = \left(\frac{d}{dt'} + \boldsymbol{\Omega}\times\right)(\mathbf{v}' + \boldsymbol{\Omega} \times \mathbf{r}), \tag{6.6}$$

or

$$\mathbf{a} = \mathbf{a}' + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) + 2\,\boldsymbol{\Omega} \times \mathbf{v}'. \tag{6.7}$$

This equation relates the apparent acceleration, $\mathbf{a} = d^2\mathbf{r}/dt^2$, of an object with position vector $\mathbf{r}$ in the non-rotating reference frame to its apparent acceleration, $\mathbf{a}' = d^2\mathbf{r}/dt'^2$, in the rotating reference frame.

Applying Newton's second law of motion in the inertial (*i.e.*, non-rotating) reference frame, we obtain

$$m\,\mathbf{a} = \mathbf{f}. \tag{6.8}$$

Here, m is the mass of our object, and $\mathbf{f}$ is the (non-fictitious) force acting on it. Note that these quantities are the *same* in both reference frames. Making use of Equation (6.7), the apparent equation of motion of our object in the rotating reference frame takes the form

$$m\,\mathbf{a}' = \mathbf{f} - m\,\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) - 2\,m\,\boldsymbol{\Omega} \times \mathbf{v}'. \tag{6.9}$$

The last two terms in the above equation are so-called "fictitious forces". Such forces are always needed to account for motion observed in non-inertial reference frames. Note that fictitious forces can always be distinguished from non-fictitious forces in Newtonian mechanics because the former have no associated reactions. Let us now investigate the two fictitious forces appearing in Equation (6.9).

6.3 Centrifugal Acceleration

Let our non-rotating inertial frame be one whose origin lies at the center of the Earth, and let our rotating frame be one whose origin is fixed with respect to some point, of latitude λ, on the Earth's surface—see Figure 6.1. The latter reference frame thus rotates with respect to the former (about an axis passing through the Earth's center) with an angular velocity vector, $\boldsymbol{\Omega}$, which points from the center of the Earth toward its north pole, and is of magnitude

$$\Omega = \frac{2\pi}{24\text{ hrs}} = 7.27 \times 10^{-5}\,\text{rad./s}. \tag{6.10}$$

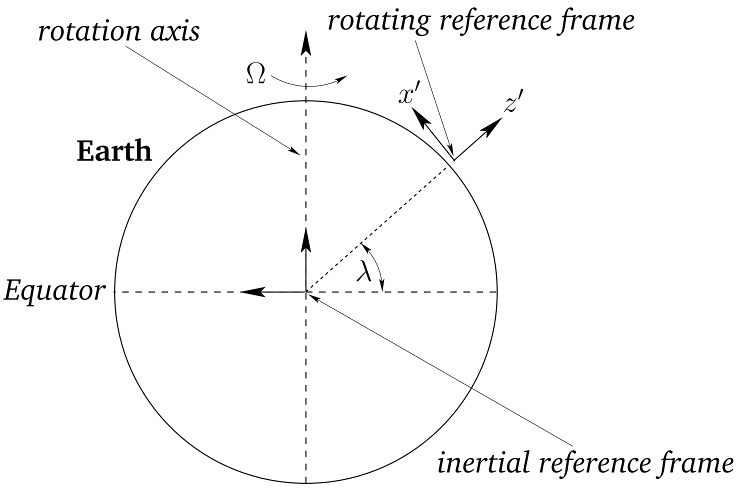


Figure 6.1: *Inertial and non-inertial reference frames.*

Consider an object which appears stationary in our rotating reference frame: *i.e.*, an object which is stationary with respect to the Earth's surface. According to Equation (6.9), the object's apparent equation of motion in the rotating frame takes the form

$$m\,\mathbf{a}' = \mathbf{f} - m\,\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}). \tag{6.11}$$

Let the non-fictitious force acting on our object be the force of gravity, $\mathbf{f} = m\,\mathbf{g}$. Here, the local gravitational acceleration, $\mathbf{g}$, points directly toward the center of the Earth. It follows, from the above, that the apparent gravitational acceleration in the rotating frame is written

$$\mathbf{g}' = \mathbf{g} - \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{R}), \tag{6.12}$$

where $\mathbf{R}$ is the displacement vector of the origin of the rotating frame (which lies on the Earth's surface) with respect to the center of the Earth. Here, we are assuming that our object is situated relatively close to the Earth's surface (*i.e.*, $\mathbf{r} \simeq \mathbf{R}$).

It can be seen, from Equation (6.12), that the apparent gravitational acceleration of a stationary object close to the Earth's surface has two components. First, the true gravitational acceleration, $\mathbf{g}$, of magnitude $g \simeq 9.8\,\mathrm{m/s^2}$, which always points directly toward the center of the Earth. Second, the so-called *centrifugal acceleration*, $-\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{R})$. This acceleration is normal to the Earth's axis of rotation, and always points directly away from this axis. The magnitude of the centrifugal acceleration is $\Omega^2\,\rho = \Omega^2\,R\,\cos\lambda$, where ρ is the perpendicular distance to the Earth's rotation axis, and $R = 6.37 \times 10^6\,\mathrm{m}$ is the Earth's radius—see Figure 6.2.

It is convenient to define Cartesian axes in the rotating reference frame such that the z'-axis points vertically upward, and x'- and y'-axes are horizontal, with the x'-axis pointing directly northward, and the y'-axis pointing directly westward—see Figure 6.1. The Cartesian components of the Earth's angular velocity are thus

$$\mathbf{\Omega} = \Omega\,(\cos\lambda,\, 0,\, \sin\lambda), \tag{6.13}$$

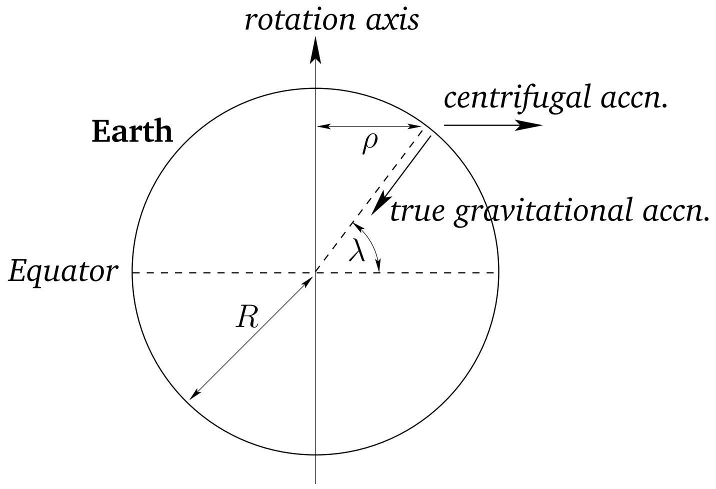


Figure 6.2: *Centrifugal acceleration.*

whilst the vectors **R** and **g** are written

$$\mathbf{R} = (0,\, 0,\, R), \tag{6.14}$$

$$\mathbf{g} = (0,\, 0,\, -g), \tag{6.15}$$

respectively. It follows that the Cartesian coordinates of the apparent gravitational acceleration, (6.12), are

$$\mathbf{g}' = \left(-\Omega^2\, R\, \cos\lambda\, \sin\lambda,\, 0,\, -g + \Omega^2\, R\, \cos^2\lambda\right). \tag{6.16}$$

The magnitude of this acceleration is approximately

$$g' \simeq g - \Omega^2\, R\, \cos^2\lambda \simeq 9.8 - 0.034\, \cos^2\lambda \;\; \mathrm{m/s^2}. \tag{6.17}$$

According to the above equation, the centrifugal acceleration causes the magnitude of the apparent gravitational acceleration on the Earth's surface to *vary* by about 0.3%, being largest at the poles, and smallest at the equator. This variation in apparent gravitational acceleration, due (ultimately) to the Earth's rotation, causes the Earth itself to *bulge* slightly at the equator (see Section 6.5), which has the effect of further intensifying the variation, since a point on the surface of the Earth at the equator is slightly further away from the Earth's center than a similar point at one of the poles (and, hence, the true gravitational acceleration is slightly weaker in the former case).

Another consequence of centrifugal acceleration is that the apparent gravitational acceleration on the Earth's surface has a *horizontal* component aligned in the north/south direction. This horizontal component ensures that the apparent gravitational acceleration *does not* point directly toward the center of the Earth. In other words, a plumb-line on the surface of the Earth does not point vertically downward, but is deflected slightly away

from a true vertical in the north/south direction. The angular deviation from true vertical can easily be calculated from Equation (6.16):

$$\theta_{dev} \simeq -\frac{\Omega^2 R}{2\,g}\,\sin(2\,\lambda) \simeq -0.1^\circ\,\sin(2\,\lambda). \tag{6.18}$$

Here, a positive angle denotes a northward deflection, and *vice versa*. Thus, the deflection is *southward* in the *northern hemisphere* (*i.e.*, $\lambda > 0$) and *northward* in the *southern hemisphere* (*i.e.*, $\lambda < 0$). The deflection is zero at the poles and at the equator, and reaches its maximum magnitude (which is very small) at middle latitudes.

6.4 Coriolis Force

We have now accounted for the first fictitious force, $-m\,\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{r})$, in Equation (6.9). Let us now investigate the second, which takes the form $-2\,m\,\boldsymbol{\Omega}\times\mathbf{v}'$, and is called the *Coriolis force*. Obviously, this force only affects objects which are *moving* in the rotating reference frame.

Consider a particle of mass m free-falling under gravity in our rotating reference frame. As before, we define Cartesian axes in the rotating frame such that the z'-axis points vertically upward, and the x'- and y'-axes are horizontal, with the x'-axis pointing directly northward, and the y'-axis pointing directly westward. It follows, from Equation (6.9), that the Cartesian equations of motion of the particle in the rotating reference frame take the form:

$$\ddot{x}' = 2\,\Omega\,\sin\lambda\,\dot{y}', \tag{6.19}$$

$$\ddot{y}' = -2\,\Omega\,\sin\lambda\,\dot{x}' + 2\,\Omega\,\cos\lambda\,\dot{z}', \tag{6.20}$$

$$\ddot{z}' = -g - 2\,\Omega\,\cos\lambda\,\dot{y}'. \tag{6.21}$$

Here, $\dot{} \equiv d/dt$, and g is the local acceleration due to gravity. In the above, we have neglected the centrifugal acceleration, for the sake of simplicity. This is reasonable, since the only effect of the centrifugal acceleration is to slightly modify the magnitude and direction of the local gravitational acceleration. We have also neglected air resistance, which is less reasonable.

Consider a particle which is dropped (at $t = 0$) from rest a height h above the Earth's surface. The following solution method exploits the fact that the Coriolis force is much smaller in magnitude that the force of gravity: hence, Ω can be treated as a *small* parameter. To lowest order (*i.e.*, neglecting Ω), the particle's vertical motion satisfies $\ddot{z}' = -g$, which can be solved, subject to the initial conditions, to give

$$z' = h - \frac{g\,t^2}{2}. \tag{6.22}$$

Substituting this expression into Equations (6.19) and (6.20), neglecting terms involving Ω^2, and solving subject to the initial conditions, we obtain $x' \simeq 0$, and

$$y' \simeq -g\,\Omega\,\cos\lambda\,\frac{t^3}{3}. \tag{6.23}$$

In other words, the particle is deflected *eastward* (*i.e.*, in the negative y'-direction). Now, the particle hits the ground when $t \simeq \sqrt{2\,h/g}$. Hence, the net eastward deflection of the particle as strikes the ground is

$$d_{\rm east} = \frac{\Omega}{3}\,\cos\lambda\,\left(\frac{8\,h^3}{g}\right)^{1/2}. \tag{6.24}$$

Note that this deflection is in the *same* direction as the Earth's rotation (*i.e.*, west to east), and is greatest at the equator, and zero at the poles. A particle dropped from a height of 100 m at the equator is deflected by about 2.2 cm.

Consider a particle launched *horizontally* with some fairly large velocity

$$\mathbf{V} = V_0\,(\cos\theta, -\sin\theta,\,0). \tag{6.25}$$

Here, θ is the *compass bearing* of the velocity vector (so north is 0°, east is 90°, *etc.*). Neglecting any vertical motion, Equations (6.19) and (6.20) yield

$$\dot{v}_{x'} \simeq -2\,\Omega\,V_0\,\sin\lambda\,\sin\theta, \tag{6.26}$$

$$\dot{v}_y' \simeq -2\,\Omega\,V_0\,\sin\lambda\,\cos\theta, \tag{6.27}$$

which can be integrated to give

$$v_{x'} \simeq V_0\,\cos\theta - 2\,\Omega\,V_0\,\sin\lambda\,\sin\theta\,t, \tag{6.28}$$

$$v_{y'} \simeq -V_0\,\sin\theta - 2\,\Omega\,V_0\,\sin\lambda\,\cos\theta\,t. \tag{6.29}$$

To lowest order in Ω, the above equations are equivalent to

$$v_{x'} \simeq V_0\,\cos(\theta + 2\,\Omega\,\sin\lambda\,t), \tag{6.30}$$

$$v_{y'} \simeq -V_0\,\sin(\theta + 2\,\Omega\,\sin\lambda\,t). \tag{6.31}$$

If follows that the Coriolis force causes the compass bearing of the particle's velocity vector to *rotate* steadily as time progresses. The rotation rate is

$$\frac{d\theta}{dt} \simeq 2\,\Omega\,\sin\lambda. \tag{6.32}$$

Hence, the rotation is *clockwise* (looking from above) in the *northern hemisphere*, and *counter-clockwise* in the *southern hemisphere*. The rotation rate is zero at the equator, and greatest at the poles.

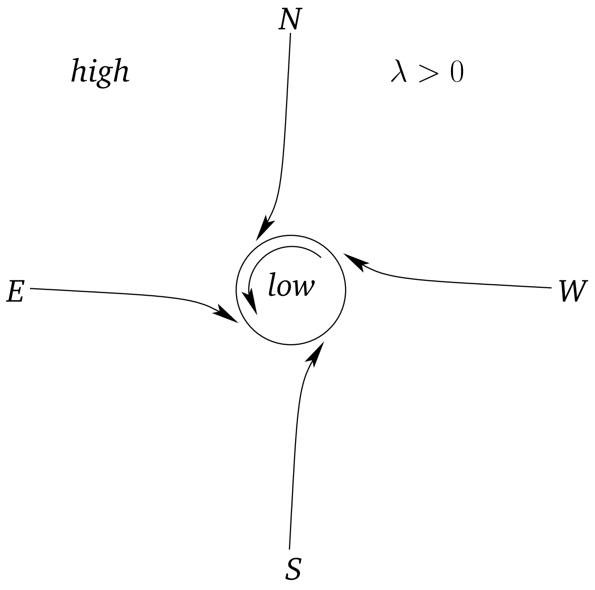


Figure 6.3: *A cyclone in the northern hemisphere.*

The Coriolis force has a significant effect on terrestrial weather patterns. Near equatorial regions, the intense heating of the Earth's surface due to the Sun results in hot air rising. In the northern hemisphere, this causes cooler air to move in a southerly direction toward the equator. The Coriolis force deflects this moving air in a clockwise sense (looking from above), resulting in the *trade winds*, which blow toward the *southwest*. In the southern hemisphere, the cooler air moves northward, and is deflected by the Coriolis force in a counter-clockwise sense, resulting in trade winds which blow toward the *northwest*.

Furthermore, as air flows from high to low pressure regions, the Coriolis force deflects the air in a clockwise/counter-clockwise manner in the northern/southern hemisphere, producing *cyclonic* rotation—see Figure 6.3. It follows that cyclonic rotation is *counter-clockwise* in the *northern hemisphere*, and *clockwise* in the *southern hemisphere*. Thus, this is the direction of rotation of tropical storms (*e.g.*, hurricanes, typhoons) in each hemisphere.

6.5 Rotational Flattening of the Earth

Consider the equilibrium configuration of a self-gravitating celestial body, composed of uniform density incompressible fluid, which is *rotating* steadily and uniformly about some fixed axis. Let the outer boundary of the body be *spheroidal*. Furthermore, let M be the body's total mass, a its mean radius, ϵ its ellipticity, and Ω its angular velocity of rotation. Finally, let the body's axis of rotation coincide with its axis of symmetry, which is assumed to run along the z-axis.

Let us transform to a non-inertial frame of reference which co-rotates with the body about the z-axis, and in which the body consequently appears to be stationary. From

Section 6.3, the problem is now analogous to that of a non-rotating body, except that the surface acceleration is written $\mathbf{g} = \mathbf{g}_g + \mathbf{g}_c$, where $\mathbf{g}_g = -\nabla\Phi(r,\theta)$ is the gravitational acceleration, and $\mathbf{g}_c$ the *centrifugal* acceleration. The latter acceleration is of magnitude $r \sin\theta\,\Omega^2$, and is everywhere directed away from the axis of rotation (see Section 6.2). The acceleration thus has components

$$\mathbf{g}_c = r\,\Omega^2\,\sin\theta\,(\sin\theta,\,\cos\theta,\,0) \tag{6.33}$$

in spherical polar coordinates. It follows that $\mathbf{g}_c = -\nabla\chi$, where

$$\chi(r,\theta) = -\frac{\Omega^2\,r^2}{2}\,\sin^2\theta = -\frac{\Omega^2\,r^2}{3}\,[1 - P_2(\cos\theta)] \tag{6.34}$$

can be thought of as a sort of centrifugal potential. Hence, the total surface acceleration is

$$\mathbf{g} = -\nabla(\Phi + \chi). \tag{6.35}$$

As before (see Section 3.6), the criterion for an equilibrium state is that the surface lie at a constant total potential, so as to eliminate tangential surface forces which cannot be balanced by internal pressure. Hence, assuming that the surface satisfies Equation (3.55), the equilibrium configuration is specified by

$$\Phi(a_\theta,\theta) + \chi(a_\theta,\theta) = c, \tag{6.36}$$

where c is a constant. It follows from Equations (3.65) and (6.34) that, to first-order in ϵ and Ω^2,

$$-\frac{G\,M}{a}\left[1 + \frac{4}{15}\,\epsilon\,P_2(\cos\theta)\right] - \frac{\Omega^2\,a^2}{3}\,[1 - P_2(\cos\theta)] \simeq c, \tag{6.37}$$

which yields

$$\epsilon = \frac{5}{4}\,\frac{\Omega^2\,a^3}{G\,M}. \tag{6.38}$$

We conclude, from the above expression, that the equilibrium configuration of a (relatively slowly) rotating self-gravitating fluid mass is an *oblate spheroid*: *i.e.*, a sphere which is slightly flattened along its axis of rotation. The degree of flattening is proportional to the square of the rotation rate. Now, from (3.55), the mean radius of the spheroid is a, the radius at the poles (*i.e.*, along the axis of rotation) is $a_p = a\,(1 - 2\,\epsilon/3)$, and the radius at the equator (*i.e.*, perpendicular to the axis of rotation) is $a_e = a\,(1 + \epsilon/3)$—see Figure 6.4. Hence, the degree of rotational flattening can be written

$$\frac{a_e - a_p}{a} = \epsilon = \frac{5}{4}\,\frac{\Omega^2\,a^3}{G\,M}. \tag{6.39}$$

Incidentally, it is reasonable to apply the above analysis to a non-fluid self-gravitating celestial body, such as the Earth, as long as the body in question is significantly larger in

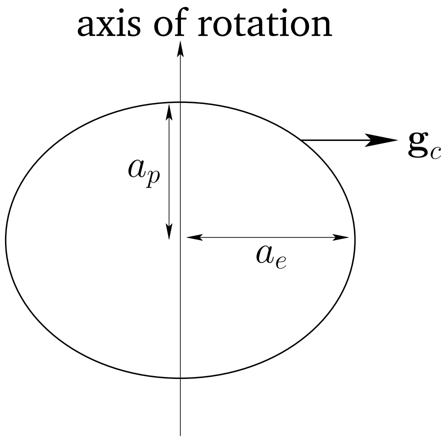


Figure 6.4: *Rotational flattening.*

diameter than about 50 km, since above this critical diameter the internal pressures that develop as a consequence of self-gravity are sufficiently large to cause most solid materials to deform like liquids (see Section 3.6). Now, for the case of the Earth, $a = 6.37 \times 10^6$ m, $\Omega = 7.27 \times 10^{-5}$ rad./s, and $M = 5.97 \times 10^{24}$ kg. Thus, we predict that

$$\epsilon = 0.00429, \tag{6.40}$$

corresponding to a difference between equatorial and polar radii of

$$\Delta a = a_e - a_p = \epsilon\, a = 27.3\,\text{km}. \tag{6.41}$$

In fact, the observed degree of flattening of the Earth is $\epsilon = 0.00335$, corresponding to a difference between equatorial and polar radii of 21.4 km. The main reason that our analysis has overestimated the degree of rotational flattening of the Earth is that it models the terrestrial interior as a uniform density incompressible fluid. In reality, the Earth's core is much denser than its crust (see Exercise 6.1).

For Jupiter, $a = 6.92 \times 10^7$ m, $\Omega = 1.76 \times 10^{-4}$ rad./s, and $M = 1.90 \times 10^{27}$ kg. Hence, we predict that

$$\epsilon = 0.101. \tag{6.42}$$

Note that this degree of flattening is much larger than that of the Earth, due to Jupiter's relatively large radius (about 10 times that of Earth), combined with its relatively short rotation period (about 0.4 days). In fact, the polar flattening of Jupiter is clearly apparent from images of this planet. The observed degree of polar flattening of Jupiter is actually $\epsilon = 0.065$. Our estimate of ϵ is probably slightly too large because Jupiter, which is mostly gaseous, has a mass distribution which is strongly concentrated at its core (see Exercise 12.1).

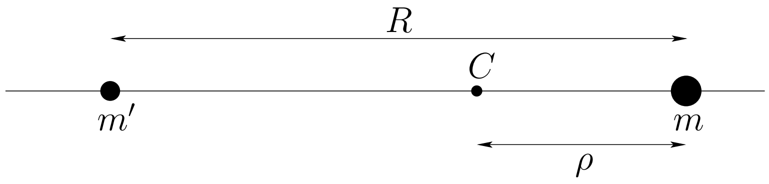


Figure 6.5: *Two orbiting masses.*

6.6 Tidal Elongation of the Earth

Consider two point masses, m and m', executing circular orbits about their common center of mass, C, with angular velocity ω. Let R be the distance between the masses, and ρ the distance between point C and mass m—see Figure 6.5. We know, from Section 4.11, that

$$\omega^2 = \frac{G\,M}{R^3}, \tag{6.43}$$

and

$$\rho = \frac{m'}{M}\,R, \tag{6.44}$$

where $M = m + m'$.

Let us transform to a non-inertial frame of reference which *rotates*, about an axis perpendicular to the orbital plane and passing through C, at the angular velocity ω. In this reference frame, both masses appear to be *stationary*. Consider mass m. In the rotating frame, this mass experiences a gravitational acceleration

$$a_g = \frac{G\,m'}{R^2} \tag{6.45}$$

directed *toward* the center of mass, and a centrifugal acceleration (see Section 6.3)

$$a_c = \omega^2\,\rho \tag{6.46}$$

directed *away from* the center of mass. However, it is easily demonstrated, using Equations (6.43) and (6.44), that

$$a_c = a_g. \tag{6.47}$$

In other words, the gravitational and centrifugal accelerations *balance*, as must be the case if mass m is to remain stationary in the rotating frame. Let us investigate how this balance is affected if the masses m and m' have *finite* spatial extents.

Let the center of the mass distribution m' lie at A, the center of the mass distribution m at B, and the center of mass at C—see Figure 6.6. We wish to calculate the centrifugal and gravitational accelerations at some point D in the vicinity of point B. It is convenient to adopt spherical coordinates, centered on point B, and aligned such that the z-axis coincides with the line BA.

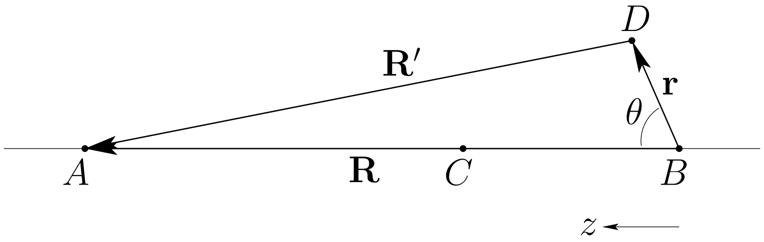


Figure 6.6: *Calculation of tidal forces.*

Let us assume that the mass distribution m is orbiting around C, but is *not* rotating about an axis passing through its center, in order to exclude rotational flattening from our analysis. If this is the case then it is easily seen that each constituent point of m executes circular motion of angular velocity ω and radius ρ—see Figure 6.7. Hence, each constituent point experiences the *same* centrifugal acceleration: *i.e.*,

$$\mathbf{g}_c = -\omega^2\,\rho\,\mathbf{e}_z. \tag{6.48}$$

It follows that

$$\mathbf{g}_c = -\nabla\chi, \tag{6.49}$$

where

$$\chi = \omega^2\,\rho\,z \tag{6.50}$$

is the centrifugal potential, and $z = r\,\cos\theta$. The centrifugal potential can also be written

$$\chi = \frac{G\,m'}{R}\frac{r}{R}\,P_1(\cos\theta). \tag{6.51}$$

The gravitational acceleration at point D due to mass m' is given by

$$\mathbf{g}_g = -\nabla\Phi', \tag{6.52}$$

where the gravitational potential takes the form

$$\Phi' = -\frac{G\,m'}{R'}. \tag{6.53}$$

Here, R' is the distance between points A and D. Note that the gravitational potential generated by the mass distribution m' is the same as that generated by an equivalent point mass at A, as long as the distribution is spherically symmetric, which we shall assume to be the case.

Now,

$$\mathbf{R}' = \mathbf{R} - \mathbf{r}, \tag{6.54}$$

where $\mathbf{R}'$ is the vector $\overrightarrow{DA}$, and $\mathbf{R}$ the vector $\overrightarrow{BA}$—see Figure 6.6. It follows that

$$R'^{-1} = \left(R^2 - 2\,\mathbf{R}\cdot\mathbf{r} + r^2\right)^{-1/2} = \left(R^2 - 2\,R\,r\,\cos\theta + r^2\right)^{-1/2}. \tag{6.55}$$

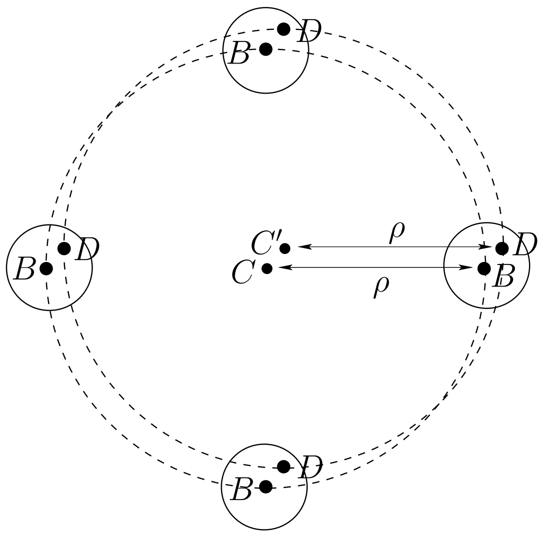


Figure 6.7: *The center* B *of the mass distribution* m *orbits about the center of mass* C *in a circle of radius* ρ*. If the mass distribution is non-rotating then a non-central point* D *must maintain a constant spatial relationship to* B*. It follows that point* D *orbits some point* C′*, which has the same spatial relationship to* C *that* D *has to* B*, in a circle of radius* ρ*.*

Expanding in powers of r/R, we obtain

$$R'^{-1} = \frac{1}{R} \sum_{n=0,\infty} \left(\frac{r}{R}\right)^n P_n(\cos\theta). \tag{6.56}$$

Hence,

$$\Phi' \simeq -\frac{G\,m'}{R}\left[1 + \frac{r}{R}\,P_1(\cos\theta) + \frac{r^2}{R^2}\,P_2(\cos\theta)\right] \tag{6.57}$$

to second-order in r/R.

Adding χ and Φ', we obtain

$$\chi + \Phi' \simeq -\frac{G\,m'}{R}\left[1 + \frac{r^2}{R^2}\,P_2(\cos\theta)\right] \tag{6.58}$$

to second-order in r/R. Note that $\chi+\Phi'$ is the potential due to the net *external* force acting on the mass distribution m. This potential is constant up to first-order in r/R, because the first-order variations in χ and Φ' cancel one another. The cancellation is a manifestation of the balance between the centrifugal and gravitational accelerations in the equivalent point mass problem discussed above. However, this balance is only exact at the *center* of the mass distribution m. Away from the center, the centrifugal acceleration remains constant, whereas the gravitational acceleration increases with increasing z. Hence, at positive z, the gravitational acceleration is larger than the centrifugal, giving rise to a net acceleration in the $+z$-direction. Likewise, at negative z, the centrifugal acceleration

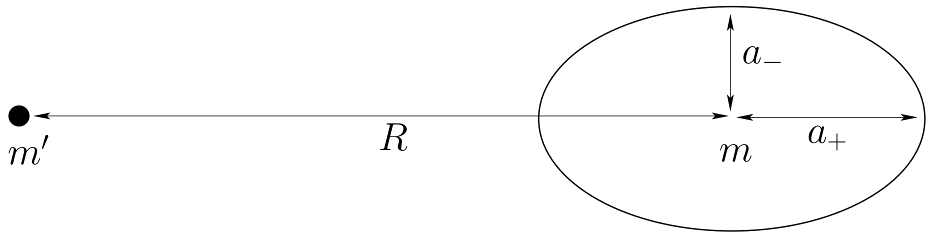


Figure 6.8: *Tidal elongation.*

is larger than the gravitational, giving rise to a net acceleration in the $-z$-direction. It follows that the mass distribution m is subject to a residual acceleration, represented by the second-order variation in Equation (6.58), which acts to *elongate* it along the z-axis. This effect is known as *tidal elongation*.

In order to calculate the tidal elongation of the mass distribution m we need to add the potential, $\chi + \Phi'$, due to the external forces, to the gravitational potential, Φ, generated by the distribution itself. Assuming that the mass distribution is spheroidal with uniform density, mass m, mean radius a, and ellipticity ϵ, it follows from Equations (3.55), (3.65), and (6.58) that the total surface potential is given by

$$\begin{aligned}\chi + \Phi' + \Phi &\simeq -\frac{G\,m}{a} - \frac{G\,m'}{R}\\ &\quad -\frac{4}{15}\,\frac{G\,m}{a}\,\epsilon\,P_2(\cos\theta) - \frac{G\,m'\,a^2}{R^3}\,P_2(\cos\theta),\end{aligned} \tag{6.59}$$

where we have treated ϵ and a/R as small quantities. As before, the condition for equilibrium is that the total potential be constant over the surface of the spheroid. Hence, we obtain

$$\epsilon = -\frac{15}{4}\,\frac{m'}{m}\left(\frac{a}{R}\right)^3 \tag{6.60}$$

as our prediction for the ellipticity induced in a self-gravitating spherical mass distribution of total mass m and radius a by a second mass, m', which is in a circular orbit of radius R about the distribution. Thus, if a_+ is the maximum radius of the distribution, and a_- the minimum radius (see Figure 6.8), then

$$\frac{a_+ - a_-}{a} = -\epsilon = \frac{15}{4}\,\frac{m'}{m}\left(\frac{a}{R}\right)^3. \tag{6.61}$$

Consider the tidal elongation of the Earth due to the Moon. In this case, we have $a = 6.37\times10^6\,\text{m}$, $R = 3.84\times10^8\,\text{m}$, $m = 5.97\times10^{24}\,\text{kg}$, and $m' = 7.35\times10^{22}\,\text{kg}$. Hence, we calculate that $-\epsilon = 2.1\times10^{-7}$, or

$$\Delta a = a_+ - a_- = -\epsilon\,a = 1.34\,\text{m}. \tag{6.62}$$

We, thus, predict that tidal forces due to the Moon cause the Earth to elongate along the axis joining its center to the Moon by about 1.3 meters. Since water is obviously more fluid

than rock (especially on relatively short time-scales) most of this elongation takes place in the oceans rather than in the underlying land. Hence, the oceans rise, relative to the land, in the region of the Earth closest to the Moon, and also in the region furthest away. Since the Earth is rotating, whilst the tidal bulge of the oceans remains relatively stationary, the Moon's tidal force causes the ocean at a given point on the Earth's surface to rise and fall, by about a meter, *twice* daily, giving rise to the phenomenon known as the *tides*.

Consider the tidal elongation of the Earth due to the Sun. In this case, we have $a = 6.37 \times 10^6$ m, $R = 1.50 \times 10^{11}$ m, $m = 5.97 \times 10^{24}$ kg, and $m' = 1.99 \times 10^{30}$ kg. Hence, we calculate that $-\epsilon = 9.6 \times 10^{-8}$, or

$$\Delta a = a_+ - a_- = -\epsilon\, a = 0.61\,\mathrm{m}. \tag{6.63}$$

Thus, the tidal elongation due to the Sun is about half that due to the Moon. It follows that the tides are particularly high when the Sun, the Earth, and the Moon lie approximately in a straight-line, so that the tidal effects of the Sun and the Moon reinforce one another. This occurs at a new moon, or at a full moon. These type of tides are called *spring tides* (note that the name has nothing to do with the season). Conversely, the tides are particularly low when the Sun, the Earth, and the Moon form a right-angle, so that the tidal effects of the Sun and the Moon partially cancel one another. These type of tides are called *neap tides*. Generally speaking, we would expect two spring tides and two neap tides per month.

In reality, the amplitude of the tides varies significantly from place to place on the Earth's surface, due to the presence of the continents, which impede the flow of the oceanic tidal bulge around the Earth. Moreover, there is an average time-lag of roughly 12 minutes between the Moon passing directly overhead (or directly below) and high tide, because of the finite inertia of the oceans. Similarly, the time-lag between a spring tide and a full moon, or a new moon, can be up to 2 days.

6.7 Tidal Torques

The fact that there is a *time-lag* between the Moon passing overhead and the occurrence of a high tide suggests the physical scenario illustrated in Figure 6.9. According to this scenario, the Moon, which is of mass m', and which is treated as a point particle, orbits the Earth (it actually orbits the center of mass of the Earth-Moon system, but this amounts to almost the same thing) in an approximately circular orbit of radius R. Moreover, the orbital angular velocity of the Moon is [see Equation (6.43)]

$$\omega = \frac{(G\,M)^{1/2}}{R^{3/2}}, \tag{6.64}$$

where $M = m + m'$ is the total mass of the Earth-Moon system. The Earth is treated as a sphere of mass m, and radius a, that rotates daily about its axis (which is approximately normal to the orbital plane of the Moon) at the angular velocity Ω. Note, incidentally, that the Earth rotates in the *same* sense that the Moon orbits, as indicated in the figure. Now, as

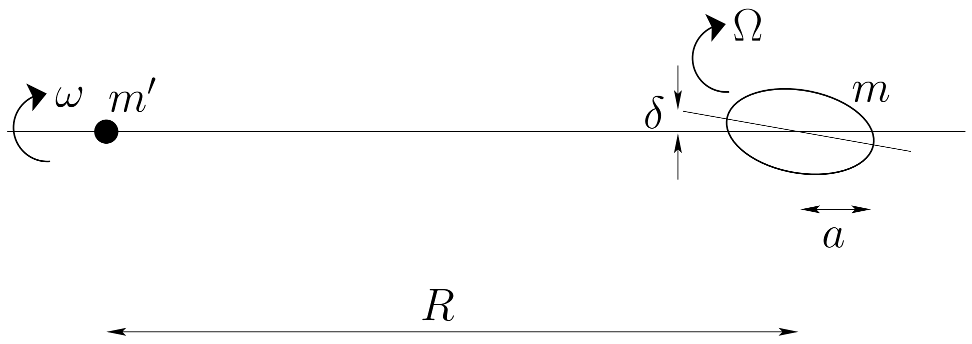


Figure 6.9: *Origin of tidal torque.*

we saw in the previous section, spatial gradients in the Moon's gravitational field produce a slight tidal elongation of the Earth. However, because of the finite inertia of the oceans, this elongation does not quite line up along the axis joining the centers of the Earth and Moon. In fact, since $\Omega > \omega$, the tidal elongation is carried *ahead* (in the sense defined by the Earth's rotation) of this axis by some small angle δ (say), as shown in the figure.

Defining a spherical coordinate system, r, θ, ϕ, whose origin is the center of the Earth, and which is orientated such that the Earth-Moon axis always corresponds to $\theta = 0$ (see Figure 6.6), the Earth's external gravitational potential is [*cf.*, Equation (3.63)]

$$\Phi(r,\theta) = -\frac{G\,m}{r} + \frac{2}{5}\,\frac{G\,m\,a^2}{r^3}\,\epsilon\left[3\,\cos^2(\theta-\delta) - 1\right], \tag{6.65}$$

where ϵ is the ellipticity induced by the tidal field of the Moon. Note that the second term on the right-hand side of the above expression is the contribution of the Earth's tidal bulge, which attains its maximum amplitude at $\theta = \delta$, rather than $\theta = 0$, because of the aforementioned misalignment between the bulge and the Earth-Moon axis. Equations (6.60) and (6.65) can be combined to give

$$\Phi(r,\theta) = -\frac{G\,m}{r} - \frac{3}{4}\,\frac{G\,m'\,a^2}{r^3}\left(\frac{a}{R}\right)^3\left[3\,\cos^2(\theta-\delta) - 1\right]. \tag{6.66}$$

Now, from (3.7), the torque that the Earth's gravitational field exerts on the Moon is

$$\tau = -m'\,\frac{\partial\Phi}{\partial\theta}\bigg|_{\theta=0,\,r=R} \simeq \frac{9}{2}\,\frac{G\,m'^2}{a}\left(\frac{a}{R}\right)^6\delta, \tag{6.67}$$

where use has been made of Equation (6.66), as well as the fact that δ is a *small* angle. Note that there is zero torque in the absence of a misalignment between the Earth's tidal bulge and the Earth-Moon axis. The torque τ acts to *increase* the Moon's orbital angular momentum. By conservation of angular momentum, an equal and opposite torque, $-\tau$, is applied to the Earth, and acts to *decrease* its rotational angular momentum. Note that if the Moon were sufficiently close to the Earth that its orbital angular velocity exceeded the Earth's rotational angular velocity (*i.e.*, if $\omega > \Omega$) then the phase-lag between the

tides and the Moon would cause the tidal bulge to fall slightly behind the Earth-Moon axis (*i.e.*, $\delta < 0$). In this case, the gravitational torque would act to reduce the Moon's orbital angular momentum, and to increase the Earth's rotational angular momentum.

The Earth's rotational equation of motion is

$$I\,\dot{\Omega} = -\tau, \tag{6.68}$$

where I is its moment of inertia. Very crudely approximating the Earth as a uniform sphere, we have $I = (2/5)\,m\,a^2$. Hence, the previous two equations can be combined to give

$$\frac{\dot{\Omega}}{\Omega} \simeq -\frac{45}{4}\,\frac{\omega^2}{\Omega}\left(\frac{m'}{m}\right)^2\left(\frac{a}{R}\right)^3\,\delta, \tag{6.69}$$

where use has been made of Equation (6.64), as well as the fact that $m \gg m'$. Now, a time-lag of 12 minutes between the Moon being overhead and the occurrence of a high-tide implies a phase-lag of $\delta \sim 0.05$ radians (*i.e.*, $\delta \sim 3^\circ$). Hence, employing the observed values $m = 5.97\times 10^{24}\,\text{kg}$, $m' = 7.35\times 10^{22}\,\text{kg}$, $a = 6.37\times 10^{6}\,\text{m}$, $R = 3.84\times 10^{8}\,\text{m}$, $\Omega = 7.27\times 10^{-5}\,\text{rad./s}$, and $\omega = 2.67\times 10^{-6}\,\text{rad./s}$, we find that

$$\frac{\dot{\Omega}}{\Omega} \simeq -3.8\times 10^{-17}\,\text{s}^{-1}. \tag{6.70}$$

It follows that, under the influence of the tidal torque, the Earth's rotation is gradually *decelerating*. Indeed, according to the above estimate, the length of a day should be increasing at the rate of about 10 milliseconds a century. (In fact, as a consequence of the severe approximations made in the our calculation, this estimate is about a factor of 5 too high.) The time-scale for the tidal torque to significantly reduce the Earth's rotational angular velocity is estimated to be

$$T \simeq \frac{\Omega}{|\dot{\Omega}|} \sim 8\times 10^{8}\,\text{years}. \tag{6.71}$$

This time-scale is comparable with the Earth's age, which is thought to be 4.5×10^{9} years. Hence, we conclude that, whilst the Earth is certainly old enough for the tidal torque to have significantly reduced its rotational angular velocity, it is plausible (especially when we take into account that our estimate for $|\dot{\Omega}|$ is 5 times too high) that it is not sufficiently old for the torque to have driven the Earth-Moon system to a final steady-state. Such a state, in which the Earth's rotational angular velocity matches the Moon's orbital angular velocity, is termed *synchronous*. In a synchronous state, the Moon would appear *stationary* to an observer on the Earth's surface, and, hence, there would be no tides (from the observer's perspective), no phase-lag, and no tidal torque.

Up to now, we have concentrated on the effect of the tidal torque on the rotation of the Earth. Let us now examine its effect on the orbit of the Moon. Now, the total angular momentum of the Earth-Moon system is

$$L \simeq \frac{2}{5}\,m\,a^2\,\Omega + m'\,R^2\,\omega, \tag{6.72}$$

where the first term on the right-hand side is the rotational angular momentum of the Earth, and the second the orbital angular momentum of the Moon. Of course, L is a conserved quantity. Moreover, ω and R are related according to Equation (6.64). It follows that

$$\frac{\dot{R}}{R} \simeq -\frac{4}{5}\frac{m}{m'}\left(\frac{a}{R}\right)^2 \frac{\dot{\Omega}}{\omega} \simeq 1.9\times 10^{-17}\,\mathrm{s}^{-1}, \tag{6.73}$$

where use has been made of (6.70). In other words, the tidal torque causes the radius of the Moon's orbit to gradually *increase*. According to the above estimate, this increase should take place at the rate of about 22 cm a year. (Again, this estimate is a factor of 5 too large.) We also have

$$\frac{\dot{\omega}}{\omega} = -\frac{3}{2}\frac{\dot{R}}{R} \sim -2.8\times 10^{-17}\,\mathrm{s}^{-1}. \tag{6.74}$$

In other words, the tidal torque produces a gradual angular *deceleration* in the Moon's orbital motion. According to the above estimate, this deceleration should take place at the rate of 150 arc seconds per century squared. (As before, this estimate is a factor of 5 too large.)

The net rate at which the tidal torques acting on the Moon and the Earth do work is

$$\dot{E} = \tau\,(\omega - \Omega). \tag{6.75}$$

Note that $\dot{E} < 0$, since $\tau > 0$ and $\Omega > \omega$. This implies that the deceleration of the Earth's rotation, and the Moon's orbital motion, that are induced by the tidal torques are necessarily associated with the *dissipation of energy*. Most of this dissipation occurs in a turbulent boundary layer at the bottoms of shallow seas such as the European shelf around the British Isles, the Patagonian shelf off Argentina, and the Bering Sea. In the absence of dissipation, there can be no steady phase-lag between the Moon and the Earth's tidal bulge, and, therefore, no tidal torques.

Of course, we would expect spatial gradients in the gravitational field of the Earth to generate a tidal bulge in the Moon. We would also expect dissipative effects to generate a phase-lag between this bulge and the Earth. This would allow the Earth to exert a gravitational torque that acts to drive the Moon toward a synchronous state in which its rotational angular velocity matches its orbital angular velocity. By analogy with the previous analysis, the de-spinning rate of the Moon is estimated to be

$$\frac{\dot{\Omega}'}{\Omega'} \simeq -\frac{45}{4}\frac{\omega^2}{\Omega'}\frac{m}{m'}\left(\frac{a'}{R}\right)^3 \delta \sim -2.3\times 10^{-13}\,\mathrm{s}^{-1}, \tag{6.76}$$

where Ω' is the Moon's rotational angular velocity, $a' = 1.74\times 10^6$ m the lunar radius, and δ the phase-lag. The above numerical estimate is made with the guesses $\Omega' = 10\,\omega$ and $\delta = 0.01$. Thus, the time required for the Moon to achieve a synchronous state is

$$T' \simeq \frac{\Omega'}{|\dot{\Omega}'|} \sim 10^5\,\text{years}. \tag{6.77}$$

This is considerably less than the age of the Moon. Hence, it is not surprising that the Moon has actually achieved a synchronous state, as evidenced by the fact that the same side of the Moon is always visible from the Earth.

6.8 Roche Radius

Consider a spherical moon of mass m and radius a which is in a circular orbit of radius R about a spherical planet of mass m' and radius a'. (Strictly speaking, the moon and the planet execute circular orbits about their common center of mass. However, if the planet is much more massive than the moon then the center of mass lies very close to the planet's center.) According to the analysis in Section 6.6, a constituent element of the moon experiences a force per unit mass, due to the gravitational field of the planet, which takes the form

$$\mathbf{g}' = -\nabla(\chi + \Phi'), \tag{6.78}$$

where

$$\chi + \Phi' = -\frac{G\,m'}{R^3}\,(z^2 - x^2/2 - y^2/2) + \text{const.} \tag{6.79}$$

Here, $(x,\ y,\ z)$ is a Cartesian coordinate system whose origin is the center of the moon, and whose z-axis always points toward the center of the planet. It follows that

$$\mathbf{g}' = \frac{2\,G\,m'}{R^3}\left(-\frac{x}{2}\,\mathbf{e}_x - \frac{y}{2}\,\mathbf{e}_y + z\,\mathbf{e}_z\right). \tag{6.80}$$

This so-called *tidal force* is generated by the spatial variation of the planet's gravitational field over the interior of the moon, and acts to elongate the moon along an axis joining its center to that of the planet, and to compress it in any direction perpendicular to this axis. Note that the magnitude of the tidal force increases strongly as the radius, R, of the moon's orbit decreases. In fact, if the tidal force becomes sufficiently strong then it can overcome the moon's self-gravity, and thereby rip the moon apart. It follows that there is a minimum radius, generally referred to as the *Roche radius*, at which a moon can orbit a planet without being destroyed by tidal forces.

Let us derive an expression for the Roche radius. Consider a small mass element at the point on the surface of the moon which lies closest to the planet, and at which the tidal force is consequently largest (*i.e.*, $x = y = 0$, $z = a$). According to Equation (6.80), the mass experiences an upward (from the moon's surface) tidal acceleration due to the gravitational attraction of the planet of the form

$$\mathbf{g}' = \frac{2\,G\,m'\,a}{R^3}\,\mathbf{e}_z. \tag{6.81}$$

The mass also experiences a downward gravitational acceleration due to the gravitational influence of the moon which is written

$$\mathbf{g} = -\frac{G\,m}{a^2}\,\mathbf{e}_z. \tag{6.82}$$

Thus, the effective surface gravity at the point in question is

$$g_{\rm eff} = \frac{G\,m}{a^2}\left(1 - 2\,\frac{m'}{m}\,\frac{a^3}{R^3}\right). \tag{6.83}$$

Note that if $R < R_c$, where

$$R_c = \left(2\,\frac{m'}{m}\right)^{1/3} a, \tag{6.84}$$

then the effective gravity is *negative*. In other words, the tidal force due to the planet is strong enough to overcome surface gravity and lift objects off the moon's surface. If this is the case, and the tensile strength of the moon is negligible, then it is fairly clear that the tidal force will start to break the moon apart. Hence, R_c is the Roche radius. Now, $m'/m = (\rho'/\rho)\,(a'/a)^3$, where ρ and ρ' are the mean mass densities of the moon and planet, respectively. Thus, the above expression for the Roche radius can also be written

$$R_c = 1.41\left(\frac{\rho'}{\rho}\right)^{1/3} a'. \tag{6.85}$$

The above calculation is somewhat inaccurate, since it fails to take into account the inevitable distortion of the moon's shape in the presence of strong tidal forces. (In fact, the calculation assumes that the moon always remains spherical.) A more accurate calculation, which treats the moon as a self-gravitating incompressible fluid, yields [1]

$$R_c = 2.44\left(\frac{\rho'}{\rho}\right)^{1/3} a'. \tag{6.86}$$

It follows that if the planet and the moon have the same mean density then the Roche radius is 2.44 times the planet's radius. Note that small orbital bodies such as rocks, or even very small moons, can survive intact within the Roche radius because they are held together by internal tensile forces rather than gravitational attraction. However, this mechanism becomes progressively less effective as the size of the body in question increases (see Section 3.6). Not surprisingly, virtually all large planetary moons occurring in the Solar System have orbital radii which exceed the relevant Roche radius, whereas virtually all planetary ring systems (which consist of myriads of small orbiting rocks) have radii which lie inside the relevant Roche radius.

6.9 Exercises

6.1. A pebble is dropped down an elevator shaft in the Empire State Building ($h = 1250$ ft, latitude $= 41^\circ$ N). Find the pebble's horizontal deflection (magnitude and direction) due to the Coriolis force at the bottom of the shaft. Neglect air resistance.

[1] *Ellipsoidal Figures of Equilibrium* S. Chandrasekhar (Yale University Press, New Haven CT, 1969).

6.2. If a bullet is fired due east, at an elevation angle α, from a point on the Earth whose latitude is $+\lambda$ show that it will strike the Earth with a lateral deflection $4\,\Omega\,v_0^3\,\sin\lambda\,\sin^2\alpha\,\cos\alpha/g^2$. Is the deflection northward or southward? Here, Ω is the Earth's angular velocity, v_0 is the bullet's initial speed, and g is the acceleration due to gravity. Neglect air resistance.

6.3. A particle is thrown vertically with initial speed v_0, reaches a maximum height, and falls back to the ground. Show that the horizontal Coriolis deflection of the particle when it returns to the ground is opposite in direction, and four times greater in magnitude, than the Coriolis deflection when it is dropped at rest from the same maximum height. Neglect air resistance.

6.4. A satellite is in a circular orbit of radius a about the Earth. Let us define a set of co-moving Cartesian coordinates, centered on the satellite, such that the x-axis always points toward the center of the Earth, the y-axis in the direction of the satellite's orbital motion, and the z-axis in the direction of the satellite's orbital angular velocity, $\boldsymbol{\omega}$. Demonstrate that the equation of motion of a small mass in orbit about the satellite are

$$\begin{aligned} \ddot{x} &= 3\,\omega^2\,x + 2\,\omega\,\dot{y}, \\ \ddot{y} &= -2\,\omega\,\dot{x}, \end{aligned}$$

assuming that $|x|/a \ll 1$ and $|y|/a \ll 1$. You may neglect the gravitational attraction between the satellite and the mass. Show that the mass executes a retrograde (*i.e.*, in the opposite sense to the satellite's orbital rotation) elliptical orbit about the satellite whose period matches that of the satellite's orbit, and whose major and minor axes are in the ratio 2 : 1, and are aligned along the y- and x-axes, respectively.

6.5. Show that

$$\epsilon = \frac{5-\alpha}{4}\,\frac{\Omega^2\,a^3}{G\,M}.$$

for a self-gravitating, rotating spheroid of ellipticity $\epsilon \ll 1$, mass M, mean radius a, and angular velocity Ω whose mass density varies as $r^{-\alpha}$ (where $\alpha < 3$). Demonstrate that the above formula matches the observed rotational flattening of the Earth when $\alpha = 1.09$ and of Jupiter when $\alpha = 1.79$.

6.6. The Moon's orbital period about the Earth is approximately 27.3 days, and is in the same direction as the Earth's axial rotation (whose period is 24 hours). Use this data to show that high tides at a given point on the Earth occur every 12 hours and 26 minutes.

6.7. Estimate the tidal elongation of the Moon due to the Earth.

7 Lagrangian Mechanics

7.1 Introduction

This chapter describes an elegant reformulation of the laws of Newtonian mechanics that is due to the French-Italian scientist Joseph Louis Lagrange. This reformulation is particularly useful for finding the equations of motion of complicated dynamical systems.

7.2 Generalized Coordinates

Let the q_i, for $i = 1, \mathcal{F}$, be a set of coordinates which uniquely specifies the instantaneous configuration of some dynamical system. Here, it is assumed that each of the q_i can vary *independently*. The q_i might be Cartesian coordinates, or polar coordinates, or angles, or some mixture of all three types of coordinate, and are, therefore, termed *generalized coordinates*. A dynamical system whose instantaneous configuration is fully specified by $\mathcal{F}$ independent generalized coordinates is said to have $\mathcal{F}$ *degrees of freedom*. For instance, the instantaneous position of a particle moving freely in three dimensions is completely specified by its three Cartesian coordinates, x, y, and z. Moreover, these coordinates are clearly independent of one another. Hence, a dynamical system consisting of a single particle moving freely in three dimensions has three degrees of freedom. If there are two freely moving particles then the system has six degrees of freedom, and so on.

Suppose that we have a dynamical system consisting of N particles moving freely in three dimensions. This is an $\mathcal{F} = 3\,N$ degree of freedom system whose instantaneous configuration can be specified by $\mathcal{F}$ Cartesian coordinates. Let us denote these coordinates the x_j, for $j = 1, \mathcal{F}$. Thus, x_1, x_2, x_3 are the Cartesian coordinates of the first particle, x_4, x_5, x_6 the Cartesian coordinates of the second particle, *etc*. Suppose that the instantaneous configuration of the system can also be specified by $\mathcal{F}$ generalized coordinates, which we shall denote the q_i, for $i = 1, \mathcal{F}$. Thus, the q_i might be the spherical coordinates of the particles. In general, we expect the x_j to be functions of the q_i. In other words,

$$x_j = x_j(q_1, q_2, \cdots, q_{\mathcal{F}}, t) \tag{7.1}$$

for $j = 1, \mathcal{F}$. Here, for the sake of generality, we have included the possibility that the functional relationship between the x_j and the q_i might depend on the time, t, explicitly. This would be the case if the dynamical system were subject to time varying constraints. For instance, a system consisting of a particle constrained to move on a surface which is itself moving. Finally, by the chain rule, the variation of the x_j due to a variation of the q_i (at constant t) is given by

$$\delta x_j = \sum_{i=1,\mathcal{F}} \frac{\partial x_j}{\partial q_i}\,\delta q_i, \tag{7.2}$$

for $j = 1, \mathcal{F}$.

7.3 Generalized Forces

The work done on the dynamical system when its Cartesian coordinates change by δx_j is simply

$$\delta W = \sum_{j=1,\mathcal{F}} f_j\, \delta x_j \tag{7.3}$$

Here, the f_j are the Cartesian components of the forces acting on the various particles making up the system. Thus, f_1, f_2, f_3 are the components of the force acting on the first particle, f_4, f_5, f_6 the components of the force acting on the second particle, *etc*. Using Equation (7.2), we can also write

$$\delta W = \sum_{j=1,\mathcal{F}} f_j \sum_{i=1,\mathcal{F}} \frac{\partial x_j}{\partial q_i}\, \delta q_i. \tag{7.4}$$

The above expression can be rearranged to give

$$\delta W = \sum_{i=1,\mathcal{F}} Q_i\, \delta q_i, \tag{7.5}$$

where

$$Q_i = \sum_{j=1,\mathcal{F}} f_j\, \frac{\partial x_j}{\partial q_i}. \tag{7.6}$$

Here, the Q_i are termed *generalized forces*. Note that a generalized force does not necessarily have the dimensions of force. However, the product $Q_i\, q_i$ must have the dimensions of work. Thus, if a particular q_i is a Cartesian coordinate then the associated Q_i is a force. Conversely, if a particular q_i is an angle then the associated Q_i is a torque.

Suppose that the dynamical system in question is *conservative*. It follows that

$$f_j = -\frac{\partial U}{\partial x_j}, \tag{7.7}$$

for $j = 1, \mathcal{F}$, where $U(x_1, x_2, \cdots, x_{\mathcal{F}}, t)$ is the system's potential energy. Hence, according to Equation (7.6),

$$Q_i = -\sum_{j=1,\mathcal{F}} \frac{\partial U}{\partial x_j}\, \frac{\partial x_j}{\partial q_i} = -\frac{\partial U}{\partial q_i}, \tag{7.8}$$

for $i = 1, \mathcal{F}$.

7.4 Lagrange's Equation

The Cartesian equations of motion of our system take the form

$$m_j\, \ddot{x}_j = f_j, \tag{7.9}$$

for $j = 1, \mathcal{F}$, where m_1, m_2, m_3 are each equal to the mass of the first particle, m_4, m_5, m_6 are each equal to the mass of the second particle, *etc.* Furthermore, the kinetic energy of the system can be written

$$K = \frac{1}{2} \sum_{j=1,\mathcal{F}} m_j\, \dot{x}_j^{\,2}. \tag{7.10}$$

Now, since $x_j = x_j(q_1, q_2, \cdots, q_{\mathcal{F}}, t)$, we can write

$$\dot{x}_j = \sum_{i=1,\mathcal{F}} \frac{\partial x_j}{\partial q_i}\, \dot{q}_i + \frac{\partial x_j}{\partial t}, \tag{7.11}$$

for $j = 1, \mathcal{F}$. Hence, it follows that $\dot{x}_j = \dot{x}_j(\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_{\mathcal{F}}, q_1, q_2, \cdots, q_{\mathcal{F}}, t)$. According to the above equation,

$$\frac{\partial \dot{x}_j}{\partial \dot{q}_i} = \frac{\partial x_j}{\partial q_i}, \tag{7.12}$$

where we are treating the $\dot{q}_i$ and the q_i as *independent* variables.

Multiplying Equation (7.12) by $\dot{x}_j$, and then differentiating with respect to time, we obtain

$$\frac{d}{dt}\left(\dot{x}_j\, \frac{\partial \dot{x}_j}{\partial \dot{q}_i}\right) = \frac{d}{dt}\left(\dot{x}_j\, \frac{\partial x_j}{\partial q_i}\right) = \ddot{x}_j\, \frac{\partial x_j}{\partial q_i} + \dot{x}_j\, \frac{d}{dt}\left(\frac{\partial x_j}{\partial q_i}\right). \tag{7.13}$$

Now,

$$\frac{d}{dt}\left(\frac{\partial x_j}{\partial q_i}\right) = \sum_{k=1,\mathcal{F}} \frac{\partial^2 x_j}{\partial q_i\, \partial q_k}\, \dot{q}_k + \frac{\partial^2 x_j}{\partial q_i\, \partial t}. \tag{7.14}$$

Furthermore,

$$\frac{1}{2}\, \frac{\partial \dot{x}_j^{\,2}}{\partial \dot{q}_i} = \dot{x}_j\, \frac{\partial \dot{x}_j}{\partial \dot{q}_i}, \tag{7.15}$$

and

$$\begin{aligned}
\frac{1}{2}\, \frac{\partial \dot{x}_j^{\,2}}{\partial q_i} = \dot{x}_j\, \frac{\partial \dot{x}_j}{\partial q_i} &= \dot{x}_j\, \frac{\partial}{\partial q_i}\left(\sum_{k=1,\mathcal{F}} \frac{\partial x_j}{\partial q_k}\, \dot{q}_k + \frac{\partial x_j}{\partial t}\right) \\
&= \dot{x}_j\left(\sum_{k=1,\mathcal{F}} \frac{\partial^2 x_j}{\partial q_i\, \partial q_k}\, \dot{q}_k + \frac{\partial^2 x_j}{\partial q_i\, \partial t}\right) \\
&= \dot{x}_j\, \frac{d}{dt}\left(\frac{\partial x_j}{\partial q_i}\right),
\end{aligned} \tag{7.16}$$

where use has been made of Equation (7.14). Thus, it follows from Equations (7.13), (7.15), and (7.16) that

$$\frac{d}{dt}\left(\frac{1}{2}\, \frac{\partial \dot{x}_j^{\,2}}{\partial \dot{q}_i}\right) = \ddot{x}_j\, \frac{\partial x_j}{\partial q_i} + \frac{1}{2}\, \frac{\partial \dot{x}_j^{\,2}}{\partial q_i}. \tag{7.17}$$

Let us take the above equation, multiply by m_j, and then sum over all j. We obtain

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{q}_i}\right) = \sum_{j=1,\mathcal{F}} f_j \frac{\partial x_j}{\partial q_i} + \frac{\partial K}{\partial q_i}, \tag{7.18}$$

where use has been made of Equations (7.9) and (7.10). Thus, it follows from Equation (7.6) that

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{q}_i}\right) = Q_i + \frac{\partial K}{\partial q_i}. \tag{7.19}$$

Finally, making use of Equation (7.8), we get

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{q}_i}\right) = -\frac{\partial U}{\partial q_i} + \frac{\partial K}{\partial q_i}. \tag{7.20}$$

It is helpful to introduce a function L, called the *Lagrangian*, which is defined as the difference between the kinetic and potential energies of the dynamical system under investigation:

$$L = K - U. \tag{7.21}$$

Since the potential energy U is clearly independent of the $\dot{q}_i$, it follows from Equation (7.20) that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \tag{7.22}$$

for $i = 1, \mathcal{F}$. This equation is known as *Lagrange's equation*.

According to the above analysis, if we can express the kinetic and potential energies of our dynamical system solely in terms of our generalized coordinates and their time derivatives then we can immediately write down the equations of motion of the system, expressed in terms of the generalized coordinates, using Lagrange's equation, (7.22). Unfortunately, this scheme only works for conservative systems. Let us now consider some examples.

Consider a particle of mass m moving in two dimensions in the central potential $U(r)$. This is clearly a two degree of freedom dynamical system. As described in Section 4.4, the particle's instantaneous position is most conveniently specified in terms of the plane polar coordinates r and θ. These are our two generalized coordinates. According to Equation (4.13), the square of the particle's velocity can be written

$$v^2 = \dot{r}^2 + (r\,\dot{\theta})^2. \tag{7.23}$$

Hence, the Lagrangian of the system takes the form

$$L = \frac{1}{2}\,m\,(\dot{r}^2 + r^2\,\dot{\theta}^2) - U(r). \tag{7.24}$$

Note that

$$\frac{\partial L}{\partial \dot{r}} = m\,\dot{r}, \qquad \frac{\partial L}{\partial r} = m\,r\,\dot{\theta}^{\,2} - \frac{dU}{dr}, \tag{7.25}$$

$$\frac{\partial L}{\partial \dot{\theta}} = m\,r^2\,\dot{\theta}, \qquad \frac{\partial L}{\partial \theta} = 0. \tag{7.26}$$

Now, Lagrange's equation (7.22) yields the equations of motion,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0, \tag{7.27}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0. \tag{7.28}$$

Hence, we obtain

$$\frac{d}{dt}(m\,\dot{r}) - m\,r\,\dot{\theta}^{\,2} + \frac{dU}{dr} = 0, \tag{7.29}$$

$$\frac{d}{dt}\left(m\,r^2\,\dot{\theta}\right) = 0, \tag{7.30}$$

or

$$\ddot{r} - r\,\dot{\theta}^{\,2} = -\frac{dV}{dr}, \tag{7.31}$$

$$r^2\,\dot{\theta} = h, \tag{7.32}$$

where $V = U/m$, and h is a constant. We recognize Equations (7.31) and (7.32) as the equations we derived in Chapter 4 for motion in a central potential. The advantage of the Lagrangian method of deriving these equations is that we avoid having to express the acceleration in terms of the generalized coordinates r and θ.

7.5 Generalized Momenta

Consider the motion of a single particle moving in one dimension. The kinetic energy is

$$K = \frac{1}{2}\,m\,\dot{x}^{\,2}, \tag{7.33}$$

where m is the mass of the particle, and x its displacement. Now, the particle's linear momentum is $p = m\,\dot{x}$. However, this can also be written

$$p = \frac{\partial K}{\partial \dot{x}} = \frac{\partial L}{\partial \dot{x}}, \tag{7.34}$$

since $L = K - U$, and the potential energy U is independent of $\dot{x}$.

Consider a dynamical system described by $\mathcal{F}$ generalized coordinates q_i, for $i = 1, \mathcal{F}$. By analogy with the above expression, we can define *generalized momenta* of the form

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{7.35}$$

for $i = 1, \mathcal{F}$. Here, p_i is sometimes called the momentum *conjugate* to the coordinate q_i. Hence, Lagrange's equation (7.22) can be written

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i}, \tag{7.36}$$

for $i = 1, \mathcal{F}$. Note that a generalized momentum does not necessarily have the dimensions of linear momentum.

Suppose that the Lagrangian L does not depend explicitly on some coordinate q_k. It follows from Equation (7.36) that

$$\frac{dp_k}{dt} = \frac{\partial L}{\partial q_k} = 0. \tag{7.37}$$

Hence,

$$p_k = \text{const.} \tag{7.38}$$

The coordinate q_k is said to be *ignorable* in this case. Thus, we conclude that the generalized momentum associated with an ignorable coordinate is a constant of the motion.

For example, the Lagrangian (7.24) for a particle moving in a central potential is independent of the angular coordinate θ. Thus, θ is an ignorable coordinate, and

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = m\, r^2\, \dot{\theta} \tag{7.39}$$

is a constant of the motion. Of course, p_θ is the angular momentum about the origin. This is conserved because a central force exerts no torque about the origin.

7.6 Exercises

7.1. A horizontal rod AB rotates with constant angular velocity ω about its midpoint O. A particle P is attached to it by equal strings AP, BP. If θ is the inclination of the plane APB to the vertical, prove that

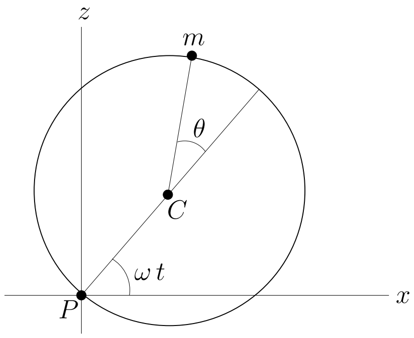
$$\frac{d^2\theta}{dt^2} - \omega^2 \sin\theta \cos\theta = -\frac{g}{l} \sin\theta,$$

where $l = OP$. Deduce the condition that the vertical position of OP should be stable.

7.2. A double pendulum consists of two simple pendula, with one pendulum suspended from the bob of the other. If the two pendula have equal lengths, l, and have bobs of equal mass, m, and if both pendula are confined to move in the same vertical plane, find Lagrange's equations of motion for the system. Use θ and ϕ—the angles the upper and lower pendulums make with the downward vertical (respectively)—as the generalized coordinates. Do not assume small angles.

7.3. Find Lagrange's equations of motion for an elastic pendulum consisting of a particle of mass m attached to an elastic string of stiffness k and unstretched length l_0. Assume that the motion takes place in a vertical plane.

7.4. A disk of mass M and radius R rolls without slipping down a plane inclined at an angle α to the horizontal. The disk has a short weightless axle of negligible radius. From this axle is suspended a simple pendulum of length $l < R$ whose bob is of mass m. Assume that the motion of the pendulum takes place in the plane of the disk. Find Lagrange's equations of motion of the system.



7.5. A vertical circular hoop of radius a is rotated in a vertical plane about a point P on its circumference at the constant angular velocity ω. A bead of mass m slides without friction on the hoop. Find the kinetic energy, the potential energy, the Lagrangian, and Largrange's equation of motion of the bead, respectively, in terms of the angular coordinate θ shown in the above diagram. Here, x is a horizontal Cartesian coordinate, z a vertical Cartesian coordinate, and C the center of the hoop.

7.6. The kinetic energy of a rotating rigid object with an axis of symmetry can be written

$$K = \frac{1}{2}\left[I_\perp\,\dot{\theta}^{\,2} + (I_\perp\,\sin^2\theta + I_\parallel\,\cos^2\theta)\,\dot{\phi}^{\,2} + 2\,I_\parallel\,\cos\theta\,\dot{\phi}\,\dot{\psi} + I_\parallel\,\dot{\psi}^{\,2}\right],$$

where $I_\parallel$ is the moment of inertia about the symmetry axis, $I_\perp$ is the moment of inertia about an axis perpendicular to the symmetry axis, and θ, ϕ, ψ are the three Euler angles. Suppose that the object is rotating freely. Find the momenta conjugate to the Euler angles. Which of these momenta are conserved? Find Lagrange's equations of motion for the system. Demonstrate that if the system is precessing steadily (which implies that θ, $\dot{\phi}$, and $\dot{\psi}$ are constants) then

$$\dot{\psi} = \left(\frac{I_\perp - I_\parallel}{I_\parallel}\right)\cos\theta\,\dot{\phi}.$$

8 Rigid Body Rotation

8.1 Introduction

This chapter examines the rotation of rigid bodies (*e.g.*, the Planets) in three dimensions.

8.2 Fundamental Equations

We can think of a rigid body as a collection of a large number of small mass elements which all maintain a fixed spatial relationship with respect to one another. Let there be N elements, and let the ith element be of mass m_i, and instantaneous position vector $\mathbf{r}_i$. The equation of motion of the ith element is written

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{j=1,N}^{j\neq i} \mathbf{f}_{ij} + \mathbf{F}_i. \tag{8.1}$$

Here, $\mathbf{f}_{ij}$ is the internal force exerted on the ith element by the jth element, and $\mathbf{F}_i$ the external force acting on the ith element. The internal forces $\mathbf{f}_{ij}$ represent the stresses which develop within the body in order to ensure that its various elements maintain a constant spatial relationship with respect to one another. Of course, $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$, by Newton's third law. The external forces represent forces which originate outside the body.

Repeating the analysis of Section 2.6, we can sum Equation (8.1) over all mass elements to obtain

$$M \frac{d^2\mathbf{r}_{cm}}{dt^2} = \mathbf{F}. \tag{8.2}$$

Here, $M = \sum_{i=1,N} m_i$ is the total mass, $\mathbf{r}_{cm}$ the position vector of the center of mass [see Equation (2.27)], and $\mathbf{F} = \sum_{i=1,N} \mathbf{F}_i$ the total external force. It can be seen that the center of mass of a rigid body moves under the action of the external forces like a point particle whose mass is identical with that of the body.

Again repeating the analysis of Section 2.6, we can sum $\mathbf{r}_i\times$ Equation (8.1) over all mass elements to obtain

$$\frac{d\mathbf{L}}{dt} = \mathbf{T}. \tag{8.3}$$

Here, $\mathbf{L} = \sum_{i=1,N} m_i\,\mathbf{r}_i \times d\mathbf{r}_i/dt$ is the total angular momentum of the body (about the origin), and $\mathbf{T} = \sum_{i=1,N} \mathbf{r}_i \times \mathbf{F}_i$ the total external torque (about the origin). Note that the above equation is only valid if the internal forces are *central* in nature. However, this is not a particularly onerous constraint. Equation (8.3) describes how the angular momentum of a rigid body evolves in time under the action of the external torques.

In the following, we shall only consider the *rotational* motion of rigid bodies, since their translational motion is similar to that of point particles [see Equation (8.2)], and, therefore, fairly straightforward in nature.

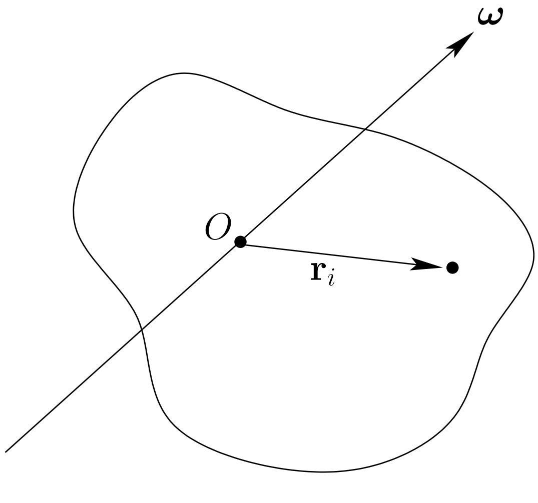


Figure 8.1: *A rigid rotating body.*

8.3 Moment of Inertia Tensor

Consider a rigid body rotating with fixed angular velocity $\boldsymbol{\omega}$ about an axis which passes through the origin—see Figure 8.1. Let $\mathbf{r}_i$ be the position vector of the ith mass element, whose mass is m_i. We expect this position vector to *precess* about the axis of rotation (which is parallel to $\boldsymbol{\omega}$) with angular velocity ω. It, therefore, follows from Equation (A.49) that

$$\frac{d\mathbf{r}_i}{dt} = \boldsymbol{\omega} \times \mathbf{r}_i. \tag{8.4}$$

Thus, the above equation specifies the velocity, $\mathbf{v}_i = d\mathbf{r}_i/dt$, of each mass element as the body rotates with fixed angular velocity $\boldsymbol{\omega}$ about an axis passing through the origin.

The total angular momentum of the body (about the origin) is written

$$\mathbf{L} = \sum_{i=1,N} m_i\, \mathbf{r}_i \times \frac{d\mathbf{r}_i}{dt} = \sum_{i=1,N} m_i\, \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i) = \sum_{i=1,N} m_i \left[r_i^2\, \boldsymbol{\omega} - (\mathbf{r}_i \cdot \boldsymbol{\omega})\, \mathbf{r}_i\right], \tag{8.5}$$

where use has been made of Equation (8.4), and some standard vector identities (see Section A.10). The above formula can be written as a matrix equation of the form

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}, \tag{8.6}$$

where

$$I_{xx} = \sum_{i=1,N} (y_i^2 + z_i^2)\, m_i = \int (y^2 + z^2)\, dm, \tag{8.7}$$

$$I_{yy} = \sum_{i=1,N} (x_i^2 + z_i^2)\, m_i = \int (x^2 + z^2)\, dm, \tag{8.8}$$

$$I_{zz} = \sum_{i=1,N} (x_i^2 + y_i^2)\, m_i = \int (x^2 + y^2)\, dm, \tag{8.9}$$

$$I_{xy} = I_{yx} = -\sum_{i=1,N} x_i\, y_i\, m_i = -\int x\, y\, dm, \tag{8.10}$$

$$I_{yz} = I_{zy} = -\sum_{i=1,N} y_i\, z_i\, m_i = -\int y\, z\, dm, \tag{8.11}$$

$$I_{xz} = I_{zx} = -\sum_{i=1,N} x_i\, z_i\, m_i = -\int x\, z\, dm. \tag{8.12}$$

Here, I_{xx} is called the *moment of inertia* about the x-axis, I_{yy} the moment of inertia about the y-axis, I_{xy} the xy *product of inertia*, I_{yz} the yz product of inertia, *etc*. The matrix of the I_{ij} values is known as the *moment of inertia tensor*.[1] Note that each component of the moment of inertia tensor can be written as either a sum over separate mass elements, or as an integral over infinitesimal mass elements. In the integrals, $dm = \rho\, dV$, where ρ is the mass density, and dV a volume element. Equation (8.6) can be written more succinctly as

$$\mathbf{L} = \tilde{\mathbf{I}}\, \boldsymbol{\omega}. \tag{8.13}$$

Here, it is understood that $\mathbf{L}$ and $\boldsymbol{\omega}$ are both *column vectors*, and $\tilde{\mathbf{I}}$ is the *matrix* of the I_{ij} values. Note that $\tilde{\mathbf{I}}$ is a *real symmetric* matrix: *i.e.*, $I_{ij}^* = I_{ij}$ and $I_{ji} = I_{ij}$.

In general, the angular momentum vector, $\mathbf{L}$, obtained from Equation (8.13), points in a different direction to the angular velocity vector, $\boldsymbol{\omega}$. In other words, $\mathbf{L}$ is generally *not parallel* to $\boldsymbol{\omega}$.

Finally, although the above results were obtained assuming a fixed angular velocity, they remain valid at each instant in time if the angular velocity varies.

8.4 Rotational Kinetic Energy

The instantaneous rotational kinetic energy of a rotating rigid body is written

$$K = \frac{1}{2} \sum_{i=1,N} m_i \left(\frac{d\mathbf{r}_i}{dt}\right)^2. \tag{8.14}$$

Making use of Equation (8.4), and some vector identities (see Section A.9), the kinetic energy takes the form

$$K = \frac{1}{2} \sum_{i=1,N} m_i\, (\boldsymbol{\omega} \times \mathbf{r}_i) \cdot (\boldsymbol{\omega} \times \mathbf{r}_i) = \frac{1}{2}\, \boldsymbol{\omega} \cdot \sum_{i=1,N} m_i\, \mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i). \tag{8.15}$$

[1] A tensor is the two-dimensional generalization of a vector. However, for present purposes, we can simply think of a tensor as another name for a matrix.

Hence, it follows from (8.5) that

$$K = \frac{1}{2}\,\boldsymbol{\omega}\cdot\mathbf{L}. \tag{8.16}$$

Making use of Equation (8.13), we can also write

$$K = \frac{1}{2}\,\boldsymbol{\omega}^{T}\,\tilde{\mathbf{I}}\,\boldsymbol{\omega}. \tag{8.17}$$

Here, $\boldsymbol{\omega}^T$ is the *row vector* of the Cartesian components ω_x, ω_y, ω_z, which is, of course, the transpose (denoted ^{T}) of the column vector $\boldsymbol{\omega}$. When written in component form, the above equation yields

$$K = \frac{1}{2}\left(I_{xx}\,\omega_x^2 + I_{yy}\,\omega_y^2 + I_{zz}\,\omega_z^2 + 2\,I_{xy}\,\omega_x\,\omega_y + 2\,I_{yz}\,\omega_y\,\omega_z + 2\,I_{xz}\,\omega_x\,\omega_z\right). \tag{8.18}$$

8.5 Principal Axes of Rotation

We have seen that the moment of inertia tensor, $\tilde{\mathbf{I}}$, defined in Section 8.3, takes the form of a *real symmetric three-dimensional* matrix. It therefore follows, from the standard matrix theory discussed in Section B.2, that the moment of inertia tensor possesses *three mutually orthogonal eigenvectors* which are associated with *three real eigenvalues*. Let the ith eigenvector (which can be normalized to be a unit vector) be denoted $\widehat{\boldsymbol{\omega}}_i$, and the ith eigenvalue λ_i. It then follows that

$$\tilde{\mathbf{I}}\,\widehat{\boldsymbol{\omega}}_i = \lambda_i\,\widehat{\boldsymbol{\omega}}_i, \tag{8.19}$$

for $i = 1, 3$.

The directions of the three mutually orthogonal unit vectors $\widehat{\boldsymbol{\omega}}_i$ define the three so-called *principal axes of rotation* of the rigid body under investigation. These axes are special because when the body rotates about one of them (*i.e.*, when $\boldsymbol{\omega}$ is parallel to one of them) the angular momentum vector $\mathbf{L}$ becomes *parallel* to the angular velocity vector $\boldsymbol{\omega}$. This can be seen from a comparison of Equation (8.13) and Equation (8.19).

Suppose that we reorient our Cartesian coordinate axes such the they coincide with the mutually orthogonal principal axes of rotation. In this new reference frame, the eigenvectors of $\tilde{\mathbf{I}}$ are the unit vectors, $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, and the eigenvalues are the moments of inertia about these axes, I_{xx}, I_{yy}, and I_{zz}, respectively. These latter quantities are referred to as the *principal moments of inertia*. Note that the products of inertia are all *zero* in the new reference frame. Hence, in this frame, the moment of inertia tensor takes the form of a *diagonal* matrix: *i.e.*,

$$\tilde{\mathbf{I}} = \begin{pmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix}. \tag{8.20}$$

Incidentally, it is easy to verify that $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are indeed the eigenvectors of the above matrix, with the eigenvalues I_{xx}, I_{yy}, and I_{zz}, respectively, and that $\mathbf{L} = \tilde{\mathbf{I}}\,\boldsymbol{\omega}$ is indeed parallel to $\boldsymbol{\omega}$ whenever $\boldsymbol{\omega}$ is directed along $\mathbf{e}_x$, $\mathbf{e}_y$, or $\mathbf{e}_z$.

When expressed in our new coordinate system, Equation (8.13) yields

$$\mathbf{L} = (I_{xx}\,\omega_x,\ I_{yy}\,\omega_y, I_{zz}\,\omega_z)\,, \tag{8.21}$$

whereas Equation (8.18) reduces to

$$K = \frac{1}{2}\left(I_{xx}\,\omega_x^{\,2} + I_{yy}\,\omega_y^{\,2} + I_{zz}\,\omega_z^{\,2}\right). \tag{8.22}$$

In conclusion, there are many great simplifications to be had by choosing a coordinate system whose axes coincide with the principal axes of rotation of the rigid body under investigation.

8.6 Euler's Equations

The fundamental equation of motion of a rotating body [see Equation (8.3)],

$$\mathbf{T} = \frac{d\mathbf{L}}{dt}, \tag{8.23}$$

is only valid in an *inertial* frame. However, we have seen that $\mathbf{L}$ is most simply expressed in a frame of reference whose axes are aligned along the principal axes of rotation of the body. Such a frame of reference *rotates* with the body, and is, therefore, *non-inertial*. Thus, it is helpful to define *two* Cartesian coordinate systems with the same origins. The first, with coordinates x, y, z, is a fixed inertial frame—let us denote this the *fixed frame*. The second, with coordinates x', y', z', co-rotates with the body in such a manner that the x'-, y'-, and z'-axes are always pointing along its principal axes of rotation—we shall refer to this as the *body frame*. Since the body frame co-rotates with the body, its instantaneous angular velocity is the same as that of the body. Hence, it follows from the analysis in Section 6.2 that

$$\frac{d\mathbf{L}}{dt} = \frac{d\mathbf{L}}{dt'} + \boldsymbol{\omega}\times\mathbf{L}. \tag{8.24}$$

Here, d/dt is the time derivative in the fixed frame, and d/dt' the time derivative in the body frame. Combining Equations (8.23) and (8.24), we obtain

$$\mathbf{T} = \frac{d\mathbf{L}}{dt'} + \boldsymbol{\omega}\times\mathbf{L}. \tag{8.25}$$

Now, in the body frame let $\mathbf{T} = (T_{x'},\ T_{y'}, T_{z'})$ and $\boldsymbol{\omega} = (\omega_{x'},\ \omega_{y'},\ \omega_{z'})$. It follows that $\mathbf{L} = (I_{x'x'}\,\omega_{x'},\ I_{y'y'}\,\omega_{y'},\ I_{z'z'}\,\omega_{z'})$, where $I_{x'x'}$, $I_{y'y'}$ and $I_{z'z'}$ are the principal moments of inertia. Hence, in the body frame, the components of Equation (8.25) yield

$$T_{x'} = I_{x'x'}\,\dot{\omega}_{x'} - (I_{y'y'} - I_{z'z'})\,\omega_{y'}\,\omega_{z'}, \tag{8.26}$$

$$T_{y'} = I_{y'y'}\,\dot{\omega}_{y'} - (I_{z'z'} - I_{x'x'})\,\omega_{z'}\,\omega_{x'}, \tag{8.27}$$

$$T_{z'} = I_{z'z'}\,\dot{\omega}_{z'} - (I_{x'x'} - I_{y'y'})\,\omega_{x'}\,\omega_{y'}, \tag{8.28}$$

where $\dot{} = d/dt'$. Here, we have made use of the fact that the moments of inertia of a rigid body are *constant* in time in the co-rotating body frame. The above equations are known as *Euler's equations*.

Consider a body that is *freely rotating*: *i.e.*, in the absence of external torques. Furthermore, let the body be *rotationally symmetric* about the z'-axis. It follows that $I_{x'x'} = I_{y'y'} = I_\perp$. Likewise, we can write $I_{z'z'} = I_\parallel$. In general, however, $I_\perp \neq I_\parallel$. Thus, Euler's equations yield

$$I_\perp \frac{d\omega_{x'}}{dt'} + (I_\parallel - I_\perp)\,\omega_{z'}\,\omega_{y'} = 0, \tag{8.29}$$

$$I_\perp \frac{d\omega_{y'}}{dt'} - (I_\parallel - I_\perp)\,\omega_{z'}\,\omega_{x'} = 0, \tag{8.30}$$

$$\frac{d\omega_{z'}}{dt'} = 0. \tag{8.31}$$

Clearly, $\omega_{z'}$ is a constant of the motion. Equation (8.29) and (8.30) can be written

$$\frac{d\omega_{x'}}{dt'} + \Omega\,\omega_{y'} = 0, \tag{8.32}$$

$$\frac{d\omega_{y'}}{dt'} - \Omega\,\omega_{x'} = 0, \tag{8.33}$$

where $\Omega = (I_\parallel/I_\perp - 1)\,\omega_{z'}$. As is easily demonstrated, the solution to the above equations is

$$\omega_{x'} = \omega_\perp \cos(\Omega\,t'), \tag{8.34}$$

$$\omega_{y'} = \omega_\perp \sin(\Omega\,t'), \tag{8.35}$$

where $\omega_\perp$ is a constant. Thus, the projection of the angular velocity vector onto the x'-y' plane has the fixed length $\omega_\perp$, and rotates steadily about the z'-axis with angular velocity Ω. It follows that the length of the angular velocity vector, $\omega = (\omega_{x'}^2 + \omega_{y'}^2 + \omega_{z'}^2)^{1/2}$, is a constant of the motion. Clearly, the angular velocity vector subtends some constant angle, α, with the z'-axis, which implies that $\omega_{z'} = \omega\,\cos\alpha$ and $\omega_\perp = \omega\,\sin\alpha$. Hence, the components of the angular velocity vector are

$$\omega_{x'} = \omega\,\sin\alpha\,\cos(\Omega\,t'), \tag{8.36}$$

$$\omega_{y'} = \omega\,\sin\alpha\,\sin(\Omega\,t'), \tag{8.37}$$

$$\omega_{z'} = \omega\,\cos\alpha, \tag{8.38}$$

where

$$\Omega = \omega\,\cos\alpha\left(\frac{I_\parallel}{I_\perp} - 1\right). \tag{8.39}$$

We conclude that, in the body frame, the angular velocity vector *precesses* about the symmetry axis (*i.e.*, the z'-axis) with the angular frequency Ω. Now, the components of the

angular momentum vector are

$$L_{x'} = I_\perp\,\omega\,\sin\alpha\,\cos(\Omega\,t'), \tag{8.40}$$

$$L_{y'} = I_\perp\,\omega\,\sin\alpha\,\sin(\Omega\,t'), \tag{8.41}$$

$$L_{z'} = I_\parallel\,\omega\,\cos\alpha. \tag{8.42}$$

Thus, in the body frame, the angular momentum vector is also of constant length, and precesses about the symmetry axis with the angular frequency Ω. Furthermore, the angular momentum vector subtends a constant angle θ with the symmetry axis, where

$$\tan\theta = \frac{I_\perp}{I_\parallel}\,\tan\alpha. \tag{8.43}$$

Note that the angular momentum vector, the angular velocity vector, and the symmetry axis all lie in the *same plane*: *i.e.*, $\mathbf{e}_{z'}\cdot\mathbf{L}\times\boldsymbol{\omega} = 0$, as can easily be verified. Moreover, the angular momentum vector lies between the angular velocity vector and the symmetry axis (*i.e.*, $\theta < \alpha$) for a flattened (or oblate) body (*i.e.*, $I_\perp < I_\parallel$), whereas the angular velocity vector lies between the angular momentum vector and the symmetry axis (*i.e.*, $\theta > \alpha$) for an elongated (or prolate) body (*i.e.*, $I_\perp > I_\parallel$).

8.7 Eulerian Angles

We have seen how we can solve Euler's equations to determine the properties of a rotating body in the co-rotating *body frame*. Let us now investigate how we can determine the same properties in the inertial *fixed frame*.

The fixed frame and the body frame share the same origin. Hence, we can transform from one to the other by means of an appropriate *rotation* of our coordinate axes. In general, if we restrict ourselves to rotations about one of the Cartesian axes, *three* successive rotations are required to transform the fixed frame into the body frame. There are, in fact, many different ways to combined three successive rotations in order to achieve this. In the following, we shall describe the most widely used method, which is due to Euler.

We start in the fixed frame, which has coordinates x, y, z, and unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$. Our first rotation is counterclockwise (looking down the axis) through an angle ϕ about the z-axis. The new frame has coordinates x'', y'', z'', and unit vectors $\mathbf{e}_{x''}$, $\mathbf{e}_{y''}$, $\mathbf{e}_{z''}$. According to Equations (A.17)–(A.19), the transformation of coordinates can be represented as follows:

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{8.44}$$

The angular velocity vector associated with ϕ has the magnitude $\dot{\phi}$, and is directed along $\mathbf{e}_z$ (*i.e.*, along the axis of rotation). Hence, we can write

$$\boldsymbol{\omega}_\phi = \dot{\phi}\,\mathbf{e}_z. \tag{8.45}$$

Clearly, $\dot{\phi}$ is the precession rate about the z-axis, as seen in the fixed frame.

The second rotation is counterclockwise (looking down the axis) through an angle θ about the x''-axis. The new frame has coordinates x''', y''', z''', and unit vectors $\mathbf{e}_{x'''}$, $\mathbf{e}_{y'''}$, $\mathbf{e}_{z'''}$. By analogy with Equation (8.44), the transformation of coordinates can be represented as follows:

$$\begin{pmatrix} x''' \\ y''' \\ z''' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}. \tag{8.46}$$

The angular velocity vector associated with θ has the magnitude $\dot{\theta}$, and is directed along $\mathbf{e}_{x''}$ (*i.e.*, along the axis of rotation). Hence, we can write

$$\boldsymbol{\omega}_\theta = \dot{\theta}\,\mathbf{e}_{x''}. \tag{8.47}$$

The third rotation is counterclockwise (looking down the axis) through an angle ψ about the z'''-axis. The new frame is the body frame, which has coordinates x', y', z', and unit vectors $\mathbf{e}_{x'}$, $\mathbf{e}_{y'}$, $\mathbf{e}_{z'}$. The transformation of coordinates can be represented as follows:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x''' \\ y''' \\ z''' \end{pmatrix}. \tag{8.48}$$

The angular velocity vector associated with ψ has the magnitude $\dot{\psi}$, and is directed along $\mathbf{e}_{z''}$ (*i.e.*, along the axis of rotation). Note that $\mathbf{e}_{z'''} = \mathbf{e}_{z'}$, since the third rotation is about $\mathbf{e}_{z'''}$. Hence, we can write

$$\boldsymbol{\omega}_\psi = \dot{\psi}\,\mathbf{e}_{z'}. \tag{8.49}$$

Clearly, $\dot{\psi}$ is *minus* the precession rate about the z'-axis, as seen in the body frame.

The full transformation between the fixed frame and the body frame is rather complicated. However, the following results can easily be verified:

$$\mathbf{e}_z = \sin\psi\,\sin\theta\,\mathbf{e}_{x'} + \cos\psi\,\sin\theta\,\mathbf{e}_{y'} + \cos\theta\,\mathbf{e}_{z'}, \tag{8.50}$$

$$\mathbf{e}_{x''} = \cos\psi\,\mathbf{e}_{x'} - \sin\psi\,\mathbf{e}_{y'}. \tag{8.51}$$

It follows from Equation (8.50) that $\mathbf{e}_z\cdot\mathbf{e}_{z'} = \cos\theta$. In other words, θ is the angle of inclination between the z- and z'-axes. Finally, since the total angular velocity can be written

$$\boldsymbol{\omega} = \boldsymbol{\omega}_\phi + \boldsymbol{\omega}_\theta + \boldsymbol{\omega}_\psi, \tag{8.52}$$

Equations (8.45), (8.47), and (8.49)–(8.51) yield

$$\omega_{x'} = \sin\psi\,\sin\theta\,\dot{\phi} + \cos\psi\,\dot{\theta}, \tag{8.53}$$

$$\omega_{y'} = \cos\psi\,\sin\theta\,\dot{\phi} - \sin\psi\,\dot{\theta}, \tag{8.54}$$

$$\omega_{z'} = \cos\theta\,\dot{\phi} + \dot{\psi}. \tag{8.55}$$

The angles ϕ, θ, and ψ are termed *Eulerian angles*. Each has a clear physical interpretation: ϕ is the angle of precession about the z-axis in the fixed frame, ψ is minus the angle of precession about the z'-axis in the body frame, and θ is the angle of inclination between the z- and z'-axes. Moreover, we can express the components of the angular velocity vector $\boldsymbol{\omega}$ in the body frame entirely in terms of the Eulerian angles, and their time derivatives [see Equations (8.53)–(8.55)].

Consider a freely rotating body that is rotationally symmetric about one axis (the z'-axis). In the absence of an external torque, the angular momentum vector $\mathbf{L}$ is a constant of the motion [see Equation (8.3)]. Let $\mathbf{L}$ point along the z-axis. In the previous section, we saw that the angular momentum vector subtends a constant angle θ with the axis of symmetry: *i.e.*, with the z'-axis. Hence, the time derivative of the Eulerian angle θ is zero. We also saw that the angular momentum vector, the axis of symmetry, and the angular velocity vector are coplanar. Consider an instant in time at which all of these vectors lie in the y'-z' plane. This implies that $\omega_{x'} = 0$. According to the previous section, the angular velocity vector subtends a constant angle α with the symmetry axis. It follows that $\omega_{y'} = \omega\, \sin\alpha$ and $\omega_{z'} = \omega\, \cos\alpha$. Equation (8.53) gives $\psi = 0$. Hence, Equation (8.54) yields

$$\omega\, \sin\alpha = \sin\theta\, \dot{\phi}. \tag{8.56}$$

This can be combined with Equation (8.43) to give

$$\dot{\phi} = \omega \left[1 + \left(\frac{I_\parallel^2}{I_\perp^2} - 1\right)\cos^2\alpha\right]^{1/2}. \tag{8.57}$$

Finally, Equation (8.55), together with (8.43) and (8.56), yields

$$\dot{\psi} = \omega\, \cos\alpha - \cos\theta\, \dot{\phi} = \omega\, \cos\alpha \left(1 - \frac{\tan\alpha}{\tan\theta}\right) = \omega\, \cos\alpha \left(1 - \frac{I_\parallel}{I_\perp}\right). \tag{8.58}$$

A comparison of the above equation with Equation (8.39) gives

$$\dot{\psi} = -\Omega. \tag{8.59}$$

Thus, as expected, $\dot{\psi}$ is minus the precession rate (of the angular momentum and angular velocity vectors) in the body frame. On the other hand, $\dot{\phi}$ is the precession rate (of the angular velocity vector and the symmetry axis) in the fixed frame. Note that $\dot{\phi}$ and Ω are quite dissimilar. For instance, Ω is negative for elongated bodies ($I_\parallel < I_\perp$) whereas $\dot{\phi}$ is positive definite. It follows that the precession is always in the same sense as L_z in the fixed frame, whereas the precession in the body frame is in the opposite sense to $L_{z'}$ for elongated bodies. We found, in the previous section, that for a *flattened* body the angular momentum vector lies between the angular velocity vector and the symmetry axis. This means that, in the fixed frame, the angular velocity vector and the symmetry axis lie on *opposite* sides of the fixed angular momentum vector. On the other hand, for an *elongated* body we found that the angular velocity vector lies between the angular momentum vector

and the symmetry axis. This means that, in the fixed frame, the angular velocity vector and the symmetry axis lie on the *same* side of the fixed angular momentum vector. (Recall that the angular momentum vector, the angular velocity vector, and the symmetry axis, are coplanar.)

8.8 Free Precession of the Earth

It is known that the Earth's axis of rotation is very slightly inclined to its symmetry axis (which passes through the two geographic poles). The angle α is approximately 0.2 seconds of an arc (which corresponds to a distance of about 6 m on the Earth's surface). It is also known that the ratio of the terrestrial moments of inertia is about $I_\parallel/I_\perp = 1.00327$, as determined from the Earth's oblateness—see Section 8.9. Hence, from (8.39), the precession rate of the angular velocity vector about the symmetry axis, as viewed in a geostationary reference frame, is

$$\Omega = 0.00327\,\omega, \tag{8.60}$$

giving a precession period of

$$T' = \frac{2\pi}{\Omega} = 305 \text{ days}. \tag{8.61}$$

(Of course, $2\pi/\omega = 1$ day.) The observed period of precession is about 440 days. The disagreement between theory and observation is attributed to the fact that the Earth is not perfectly rigid. The Earth's symmetry axis subtends an angle $\theta \simeq \alpha = 0.2''$ [see (8.43)] with its angular momentum vector, but lies on the opposite side of this vector to the angular velocity vector. This implies that, as viewed from space, the Earth's angular velocity vector is almost parallel to its fixed angular momentum vector, whereas its symmetry axis subtends an angle of 0.2" with both vectors, and precesses about them. The (theoretical) precession rate of the Earth's symmetry axis, as seen from space, is given by Equation (8.57):

$$\dot{\phi} = 1.00327\,\omega. \tag{8.62}$$

The associated precession period is

$$T = \frac{2\pi}{\dot{\phi}} = 0.997 \text{ days}. \tag{8.63}$$

The free precession of the Earth's symmetry axis in space, which is known as the *Chandler wobble*, since it was discovered by the American astronomer Seth Chandler in 1891, is superimposed on a much slower forced precession, with a period of about 26,000 years, caused by the small gravitational torque exerted on the Earth by the Sun and the Moon, as a consequence of the Earth's slight oblateness—see Section 8.10.

8.9 McCullough's Formula

According to Equations (3.57) and (3.62), if the Earth is modeled as spheroid of uniform density γ then its ellipticity is given by

$$\epsilon = -\int r^2\,\gamma\,P_2(\cos\theta)\,d^3\mathbf{r}\Big/\,I_0 = -\frac{1}{2}\int r^2\,\gamma\,\left(3\,\cos^2\theta - 1\right)\,d^3\mathbf{r}\Big/\,I_0, \tag{8.64}$$

where the integral is over the whole volume of the Earth, and $I_0 = (2/5)\,M\,a^2$ would be the Earth's moment of inertia were it exactly spherical. Now, the Earth's moment of inertia about its axis of rotation is given by

$$I_\parallel = \int (x^2 + y^2)\,\gamma\,d^3\mathbf{r} = \int r^2\,\gamma\,(1 - \cos^2\theta)\,d^3\mathbf{r}. \tag{8.65}$$

Here, use has been made of Equations (3.24)–(3.26). Likewise, the Earth's moment of inertia about an axis perpendicular to its axis of rotation (and passing through the Earth's center) is

$$\begin{aligned} I_\perp &= \int (y^2 + z^2)\,\gamma\,d^3\mathbf{r} = \int r^2\,\gamma\,(\sin^2\theta\,\sin^2\,\phi + \cos^2\theta)\,d^3\mathbf{r} \\ &= \int r^2\,\gamma\,\left(\frac{1}{2}\,\sin^2\theta + \cos^2\theta\right)\,d^3\mathbf{r} = \frac{1}{2}\int r^2\,\gamma\,(1 + \cos^2\theta)\,d^3\mathbf{r}, \end{aligned} \tag{8.66}$$

since the average of $\sin^2\phi$ is $1/2$ for an axisymmetric mass distribution. It follows from the above three equations that

$$\epsilon = \frac{I_\parallel - I_\perp}{I_0} \simeq \frac{I_\parallel - I_\perp}{I_\parallel}. \tag{8.67}$$

This result, which is known as *McCullough's formula*, demonstrates that the Earth's ellipticity is directly related to the difference between its principle moments of inertia. It turns out that McCullough's formula holds for *any* axially symmetric mass distribution, and not just a spheroidal distribution with uniform density. Finally, McCullough's formula can be combined with Equation (3.63) to give

$$\Phi(r,\theta) = -\frac{G\,M}{r} + \frac{G\,(I_\parallel - I_\perp)}{r^3}\,P_2(\cos\theta). \tag{8.68}$$

This is the general expression for the gravitational potential generated outside an axially symmetric mass distribution. The first term on the right-hand side is the *monopole* gravitational field which would be generated if all of the mass in the distribution were concentrated at its center of mass, whereas the second term is the *quadrupole* field generated by any deviation from spherical symmetry in the distribution.

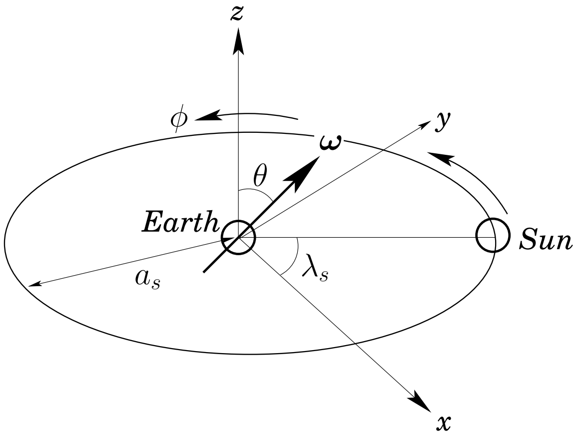


Figure 8.2: *The Earth-Sun system.*

8.10 Forced Precession and Nutation of the Earth

Consider the Earth-Sun system—see Figure 8.2. From a geocentric viewpoint, the Sun orbits the Earth counter-clockwise (looking from the north), once per year, in an approximately circular orbit of radius $a_s = 1.50 \times 10^{11}$ m. In astronomy, the plane of the Sun's apparent orbit relative to the Earth is known as the *ecliptic plane*. Let us define *non-rotating* Cartesian coordinates, centered on the Earth, which are such that the x- and y-axes lie in the ecliptic plane, and the z-axis is normal to this plane (in the sense that the Earth's north pole lies at positive z). It follows that the z-axis is directed toward a point in the sky (located in the constellation Draco) known as the *north ecliptic pole*. In the following, we shall treat the $Oxyz$ system as inertial. This is a reasonable approximation because the orbital acceleration of the Earth is much smaller than the acceleration due to its diurnal rotation. It is convenient to parameterize the instantaneous position of the Sun in terms of a counter-clockwise (looking from the north) azimuthal angle λ_s that is zero on the positive x-axis—see Figure 8.2.

Let $\boldsymbol{\omega}$ be the Earth's angular velocity vector due to its daily rotation. This vector makes an angle θ with the z-axis, where $\theta = 23.44^\circ$ is the mean inclination of the ecliptic to the Earth's equatorial plane. Suppose that the projection of $\boldsymbol{\omega}$ onto the ecliptic plane subtends an angle ϕ with the y-axis, where ϕ is measured in a counter-clockwise (looking from the north) sense—see Figure 8.2. The orientation of the Earth's axis of rotation (which is, of course, parallel to $\boldsymbol{\omega}$) is thus determined by the two angles θ and ϕ. Note, however, that these two angles are also *Euler angles*, in the sense given in Section 8.7. Let us examine the Earth-Sun system at an instant in time, $t = 0$, when $\phi = 0$: *i.e.*, when $\boldsymbol{\omega}$ lies in the y-z plane. At this particular instant, the x-axis points towards the so-called *vernal equinox*, which is defined as the point in the sky where the ecliptic plane crosses the projection of the Earth's equator (*i.e.*, the plane normal to $\boldsymbol{\omega}$) from south to north. A counter-clockwise (looking from the north) angle in the ecliptic plane that is zero at the vernal equinox is generally known as an *ecliptic longitude*. Thus, λ_s is the Sun's ecliptic longitude.

According to Equation (8.68), the potential energy of the Earth-Sun system is written

$$U = M_s\,\Phi = -\frac{G\,M_s\,M}{a_s} + \frac{G\,M_s\,(I_\parallel - I_\perp)}{a_s^3}\,P_2[\cos(\gamma_s)], \tag{8.69}$$

where M_s is the mass of the Sun, M the mass of the Earth, $I_\parallel$ the Earth's moment of inertia about its axis of rotation, $I_\perp$ the Earth's moment of inertia about an axis lying in its equatorial plane, and $P_2(x) = (1/2)\,(3\,x^2 - 1)$. Furthermore, γ_s is the angle subtended between $\boldsymbol{\omega}$ and $\mathbf{r}_s$, where $\mathbf{r}_s$ is the position vector of the Sun relative to the Earth.

It is easily demonstrated that (with $\phi = 0$)

$$\boldsymbol{\omega} = \omega\,(0,\,\sin\theta,\,\cos\theta), \tag{8.70}$$

and

$$\mathbf{r}_s = a_s\,(\cos\lambda_s,\,\sin\lambda_s,\,0). \tag{8.71}$$

Hence,

$$\cos\gamma_s = \frac{\boldsymbol{\omega}\cdot\mathbf{r}_s}{|\boldsymbol{\omega}|\,|\mathbf{r}_s|} = \sin\theta\,\sin\lambda_s, \tag{8.72}$$

giving

$$U = -\frac{G\,M_s\,M}{a_s} + \frac{G\,M_s\,(I_\parallel - I_\perp)}{2\,a_s^3}\,(3\,\sin^2\theta\,\sin^2\lambda_s - 1). \tag{8.73}$$

Now, we are primarily interested in the motion of the Earth's axis of rotation over time-scales that are *much longer* than a year, so we can average the above expression over the Sun's orbit to give

$$U = -\frac{G\,M_s\,M}{a_s} + \frac{G\,M_s\,(I_\parallel - I_\perp)}{2\,a_s^3}\left(\frac{3}{2}\,\sin^2\theta - 1\right) \tag{8.74}$$

(since the average of $\sin^2\lambda_s$ over a year is $1/2$). Thus, we obtain

$$U = U_0 - \epsilon\,\alpha_s\,\cos(2\,\theta), \tag{8.75}$$

where U_0 is a constant, and

$$\alpha_s = \frac{3}{8}\,I_\parallel\,n_s^2. \tag{8.76}$$

Here,

$$\epsilon = \frac{I_\parallel - I_\perp}{I_\parallel} = 0.00335 \tag{8.77}$$

is the Earth's ellipticity, and

$$n_s = \frac{d\lambda_s}{dt} = \left(\frac{G\,M_s}{a_s^3}\right)^{1/2} \tag{8.78}$$

the Sun's apparent orbital angular velocity.

The rotational kinetic energy of the Earth can be written (see Section 8.4)

$$K = \frac{1}{2}\left(I_\perp\,\omega_{x'}^2 + I_\perp\,\omega_{y'}^2 + I_\parallel\,\omega_{z'}^2\right), \tag{8.79}$$

which reduces to

$$K = \frac{1}{2}\left(I_\perp\,\dot{\theta}^{\,2} + I_\perp\,\sin^2\theta\,\dot{\phi}^{\,2} + I_\parallel\,\omega^2\right) \tag{8.80}$$

with the aid of Equations (8.53)–(8.55). Here,

$$\omega = \cos\theta\,\dot{\phi} + \dot{\psi}, \tag{8.81}$$

and ψ is the third Euler angle. Hence, the Earth's Lagrangian takes the form

$$\mathcal{L} = K - U = \frac{1}{2}\left(I_\perp\,\dot{\theta}^2 + I_\perp\,\sin^2\theta\,\dot{\phi}^2 + I_\parallel\,\omega^2\right) + \epsilon\,\alpha_s\,\cos(2\,\theta) \tag{8.82}$$

where any constant terms have been neglected. Note that the Lagrangian does not depend explicitly on the angular coordinate ψ. It follows that the conjugate momentum is a constant of the motion (see Section 7.5). In other words,

$$p_\psi = \frac{\partial\mathcal{L}}{\partial\dot{\psi}} = I_\parallel\,\omega \tag{8.83}$$

is a constant of the motion, implying that ω is also a constant of the motion. Note that ω is effectively the Earth's angular velocity of rotation about its axis (since $|\omega_{x'}|, |\omega_{y'}| \ll \omega_{z'} = \omega$, which follows because $|\dot{\phi}|, |\dot{\theta}| \ll \dot{\psi}$). Another equation of motion that can be derived from the Lagrangian is

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\right) - \frac{\partial\mathcal{L}}{\partial\theta} = 0, \tag{8.84}$$

which reduces to

$$I_\perp\,\ddot{\theta} - \frac{\partial\mathcal{L}}{\partial\theta} = 0. \tag{8.85}$$

Consider *steady precession* of the Earth's rotational axis, which is characterized by $\dot{\theta} = 0$, with both $\dot{\phi}$ and $\dot{\psi}$ constant. It follows, from the above equation, that such motion must satisfy the constraint

$$\frac{\partial\mathcal{L}}{\partial\theta} = 0. \tag{8.86}$$

Thus, we obtain

$$\frac{1}{2}\,I_\perp\,\sin(2\,\theta)\,\dot{\phi}^{\,2} - I_\parallel\,\sin\theta\,\omega\,\dot{\phi} - 2\,\epsilon\,\alpha_s\,\sin(2\,\theta) = 0, \tag{8.87}$$

where use has been made of Equations (8.81) and (8.82). Now, as can easily be verified after the fact, $|\dot{\phi}| \ll \omega$, so the above equation reduces to

$$\dot{\phi} \simeq -\frac{4\,\epsilon\,\alpha_s\,\cos\theta}{I_\parallel\,\omega} = \Omega_\phi, \tag{8.88}$$

which can be integrated to give

$$\phi \simeq -\Omega_\phi\, t, \tag{8.89}$$

where

$$\Omega_\phi = \frac{3}{2}\,\frac{\epsilon\, n_s^2}{\omega}\,\cos\theta, \tag{8.90}$$

and use has been made of Equation (8.76). According to the above expression, the mutual interaction between the Sun and the quadrupole gravitational field generated by the Earth's slight oblateness causes the Earth's axis of rotation to *precess* steadily about the normal to the ecliptic plane at the rate $-\Omega_\phi$. The fact that $-\Omega_\phi$ is negative implies that the precession is in the *opposite* direction to the direction of the Earth's rotation and the Sun's apparent orbit about the Earth. Incidentally, the interaction causes a precession of the Earth's rotational axis, rather than the plane of the Sun's orbit, because the Earth's axial moment of inertia is much less than the Sun's orbital moment of inertia. The precession period in years is given by

$$T_\phi(\text{yr}) = \frac{n_s}{\Omega_\phi} = \frac{2\,T_s(\text{day})}{3\,\epsilon\,\cos\theta}, \tag{8.91}$$

where $T_s(\text{day}) = \omega/n_s = 365.24$ is the Sun's orbital period in days. Thus, given that $\epsilon = 0.00335$ and $\theta = 23.44^\circ$, we obtain

$$T_\phi \simeq 79{,}200 \text{ years}. \tag{8.92}$$

Unfortunately, the observed precession period of the Earth's axis of rotation about the normal to the ecliptic plane is approximately 25,800 years, so something is clearly missing from our model. It turns out that the missing factor is the influence of the *Moon*.

Using analogous arguments to those given above, the potential energy of the Earth-Moon system can be written

$$U = -\frac{G\,M_m\,M}{a_m} + \frac{G\,M_m\,(I_\parallel - I_\perp)}{a_m^3}\,P_2[\cos(\gamma_m)], \tag{8.93}$$

where M_m is the lunar mass, and a_m the radius of the Moon's (approximately circular) orbit. Furthermore, γ_m is the angle subtended between $\boldsymbol{\omega}$ and $\mathbf{r}_m$, where

$$\boldsymbol{\omega} = \omega\,(-\sin\theta\,\sin\phi,\, \sin\theta\,\cos\phi,\, \cos\theta) \tag{8.94}$$

is the Earth's angular velocity vector, and $\mathbf{r}_m$ is the position vector of the Moon relative to the Earth. Here, for the moment, we have retained the ϕ dependence in our expression for $\boldsymbol{\omega}$ (since we shall presently differentiate by ϕ, before setting $\phi = 0$). Now, the Moon's orbital plane is actually slightly inclined to the ecliptic plane, the angle of inclination being $\iota_m = 5.16^\circ$. Hence, we can write

$$\mathbf{r}_m \simeq a_m\,(\cos\lambda_m,\, \sin\lambda_m,\, \iota_m\,\sin(\lambda_m - \varpi_n)), \tag{8.95}$$

to first order in ι_m, where λ_m is the Moon's ecliptic longitude, and ϖ_n is the ecliptic longitude of the lunar *ascending node*, which is defined as the point on the lunar orbit where

the Moon crosses the ecliptic plane from south to north. Of course, λ_m increases at the rate n_m, where

$$n_m = \frac{d\lambda_m}{dt} \simeq \left(\frac{G\,M}{a_m^3}\right)^{1/2} \tag{8.96}$$

is the Moon's orbital angular velocity. It turns out that the lunar ascending node precesses steadily, in the opposite direction to the Moon's orbital rotation, in such a manner that it completes a full circuit every 18.6 years. This precession is caused by the perturbing influence of the Sun—see Chapter 10. It follows that

$$\frac{d\varpi_n}{dt} = -\Omega_n, \tag{8.97}$$

where $2\pi/\Omega_n = 18.6$ years. Now, from (8.94) and (8.95),

$$\cos\gamma_m = \frac{\boldsymbol{\omega}\cdot\mathbf{r}_m}{|\boldsymbol{\omega}|\,|\mathbf{r}_m|} = \sin\theta\,\sin(\lambda_m - \phi) + \iota_m\,\cos\theta\,\sin(\lambda_m - \varpi_n), \tag{8.98}$$

so (8.93) yields

$$\begin{aligned} U \;\simeq\; & -\frac{G\,M_m\,M}{a_m} + \frac{G\,M_m\,(I_\parallel - I_\perp)}{2\,a_m^3}\left[3\,\sin^2\theta\,\sin^2(\lambda_m - \phi)\right. \\ & \left.+3\,\iota_m\,\sin(2\,\theta)\,\sin(\lambda_m - \phi)\,\sin(\lambda_m - \varpi_n) - 1\right] \end{aligned} \tag{8.99}$$

to first order in ι_m. Given that we are interested in the motion of the Earth's axis of rotation on time-scales that are much longer than a month, we can average the above expression over the Moon's orbit to give

$$U \simeq U_0' - \epsilon\,\alpha_m\,\cos(2\,\theta) + \epsilon\,\beta_m\,\sin(2\,\theta)\,\cos(\varpi_n - \phi), \tag{8.100}$$

[since the average of $\sin^2(\lambda_m - \phi)$ over a month is $1/2$, whereas that of $\sin(\lambda_m - \phi)\,\sin(\lambda_m - \varpi_m)$ is $(1/2)\,\cos(\varpi_m - \phi)$]. Here, U_0' is a constant,

$$\alpha_m = \frac{3}{8}\,I_\parallel\,\mu_m\,n_m^2, \tag{8.101}$$

$$\beta_m = \frac{3}{4}\,I_\parallel\,\iota_m\,\mu_m\,n_m^2, \tag{8.102}$$

and

$$\mu_m = \frac{M_m}{M} = 0.0123 \tag{8.103}$$

is the ratio of the lunar to the terrestrial mass. Now, gravity is a superposable force, so the total potential energy of the Earth-Moon-Sun system is the sum of Equations (8.75) and (8.100). In other words,

$$U = U_0'' - \epsilon\,\alpha\,\cos(2\,\theta) + \epsilon\,\beta_m\,\sin(2\,\theta)\,\cos(\varpi_n - \phi), \tag{8.104}$$

where U_0'' is a constant, and

$$\alpha = \alpha_s + \alpha_m. \tag{8.105}$$

Finally, making use of (8.80), the Lagrangian of the Earth is written

$$\mathcal{L} = \frac{1}{2}\left(I_\perp\,\dot{\theta}^2 + I_\perp\,\sin^2\theta\,\dot{\phi}^2 + I_\parallel\,\omega^2\right) + \epsilon\,\alpha\,\cos(2\,\theta) - \epsilon\,\beta_m\,\sin(2\,\theta)\,\cos(\varpi_n - \phi), \tag{8.106}$$

where any constant terms have been neglected. Recall that ω is given by (8.81), and is a constant of the motion.

Two equations of motion that can immediately be derived from the above Lagrangian are

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) - \frac{\partial \mathcal{L}}{\partial \theta} = 0, \tag{8.107}$$

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{8.108}$$

(The third equation, involving ψ, merely confirms that ω is a constant of the motion.) The above two equations yield

$$0 = I_\perp\,\ddot{\theta} - \frac{1}{2}\,I_\perp\,\sin(2\,\theta)\,\dot{\phi}^{\,2} + I_\parallel\,\sin\theta\,\omega\,\dot{\phi} + 2\,\epsilon\,\alpha\,\sin(2\,\theta) + 2\,\epsilon\,\beta_m\,\cos(2\,\theta)\,\cos(\varpi_n - \phi), \tag{8.109}$$

$$0 = \frac{d}{dt}\left(I_\perp\,\sin^2\theta\,\dot{\phi} + I_\parallel\,\cos\theta\,\omega\right) + \epsilon\,\beta_m\,\sin(2\,\theta)\,\sin(\varpi_n - \phi), \tag{8.110}$$

respectively. Let

$$\theta(t) = \theta_0 + \epsilon\,\theta_1(t), \tag{8.111}$$

$$\phi(t) = \epsilon\,\phi_1(t), \tag{8.112}$$

where $\theta_0 = 23.44^\circ$ is the mean inclination of the ecliptic to the Earth's equatorial plane. To first order in ϵ, Equations (8.109) and (8.110) reduce to

$$0 \simeq I_\perp\,\ddot{\theta}_1 + I_\parallel\,\sin\theta_0\,\omega\,\dot{\phi}_1 + 2\,\alpha\,\sin(2\,\theta_0) + 2\,\beta_m\,\cos(2\,\theta_0)\,\cos(\Omega_n\,t), \tag{8.113}$$

$$0 \simeq I_\perp\,\sin^2\theta_0\,\ddot{\phi}_1 - I_\parallel\,\sin\theta_0\,\omega\,\dot{\theta}_1 - \beta_m\,\sin(2\,\theta_0)\,\sin(\Omega_n\,t), \tag{8.114}$$

respectively, where use has been made of Equation (8.97). However, as can easily be verified after the fact, $d/dt \ll \omega$, so we obtain

$$\dot{\phi}_1 \simeq -\frac{4\,\alpha\,\cos\theta_0}{I_\parallel\,\omega} - \frac{2\,\beta_m\,\cos(2\,\theta_0)}{I_\parallel\,\omega\,\sin\theta_0}\,\cos(\Omega_n\,t), \tag{8.115}$$

$$\dot{\theta}_1 \simeq -\frac{2\,\beta_m\,\cos\theta_0}{I_\parallel\,\omega}\,\sin(\Omega_n\,t). \tag{8.116}$$

The above equations can be integrated, and then combined with Equations (8.111) and (8.112), to give

$$\phi(t) = -\Omega_\phi\, t - \delta\phi\, \sin(\Omega_n\, t), \tag{8.117}$$

$$\theta(t) = \theta_0 + \delta\theta\, \cos(\Omega_n\, t), \tag{8.118}$$

where

$$\Omega_\phi = \frac{3}{2}\,\frac{\epsilon\,(n_s^2 + \mu_m\, n_m^2)}{\omega}\,\cos\theta_0, \tag{8.119}$$

$$\delta\phi = \frac{3}{2}\,\frac{\epsilon\,\iota_m\,\mu_m\, n_m^2}{\omega\,\Omega_n}\,\frac{\cos(2\,\theta_0)}{\sin\theta_0}, \tag{8.120}$$

$$\delta\theta = \frac{3}{2}\,\frac{\epsilon\,\iota_m\,\mu_m\, n_m^2}{\omega\,\Omega_n}\,\cos\theta_0. \tag{8.121}$$

Incidentally, in the above, we have assumed that the lunar ascending node coincides with the vernal equinox at time $t = 0$ (*i.e.*, $\varpi_m = 0$ at $t = 0$), in accordance with our previous assumption that $\phi = 0$ at $t = 0$.

According to Equation (8.117), the combined gravitational interaction of the Sun and the Moon with the quadrupole field generated by the Earth's slight oblateness causes the Earth's axis of rotation to precess steadily about the normal to the ecliptic plane at the rate $-\Omega_\phi$. As before, the negative sign indicates that the precession is in the opposite direction to the (apparent) orbital motion of the sun and moon. The period of the precession in years is given by

$$T_\phi(\text{yr}) = \frac{n_s}{\Omega_\phi} = \frac{2\,T_s(\text{day})}{\epsilon\,(1 + \mu_m/[T_m(\text{yr})]^2)\,\cos\theta_0}, \tag{8.122}$$

where $T_m(\text{yr}) = n_s/n_m = 0.081$ is the Moon's (synodic) orbital period in years. Given that $\epsilon = 0.00335$, $\theta_0 = 23.44^\circ$, $T_s(\text{day}) = 365.24$, and $\mu_m = 0.0123$, we obtain

$$T_\phi \simeq 27{,}600\,\text{years}. \tag{8.123}$$

This prediction is fairly close to the observed precession period of 25, 800 years. The main reason that our estimate is slightly inaccurate is because we have neglected to take into account the small eccentricities of the Earth's orbit around the Sun, and the Moon's orbit around the Earth.

The point in the sky toward which the Earth's axis of rotation points is known as the *north celestial pole*. Currently, this point lies within about a degree of the fairly bright star *Polaris*, which is consequently sometimes known as the *north star* or the *pole star*. It follows that Polaris appears to be almost stationary in the sky, always lying *due north*, and can thus be used for navigational purposes. Indeed, mariners have relied on the north star for many hundreds of years to determine direction at sea. Unfortunately, because of the precession of the Earth's axis of rotation, the north celestial pole is not a fixed point in the sky, but instead traces out a circle, of angular radius 23.44°, about the north ecliptic pole,

with a period of 25,800 years. Hence, a few thousand years from now, the north celestial pole will no longer coincide with Polaris, and there will be no convenient way of telling direction from the stars.

The projection of the ecliptic plane onto the sky is called the *ecliptic*, and coincides with the apparent path of the Sun against the backdrop of the stars. Furthermore, the projection of the Earth's equator onto the sky is known as the *celestial equator*. As has been previously mentioned, the ecliptic is inclined at 23.44° to the celestial equator. The two points in the sky at which the ecliptic crosses the celestial equator are called the *equinoxes*, since night and day are equally long when the Sun lies at these points. Thus, the Sun reaches the *vernal equinox* on about March 21st, and this traditionally marks the beginning of spring. Likewise, the Sun reaches the *autumn equinox* on about September 22nd, and this traditionally marks the beginning of autumn. However, the precession of the Earth's axis of rotation causes the celestial equator (which is always normal to this axis) to precess in the sky, and thus also causes the equinoxes to precess along the ecliptic. This effect is known as the *precession of the equinoxes*. The precession is in the opposite direction to the Sun's apparent motion around the ecliptic, and is of magnitude 1.4° per century. Amazingly, this miniscule effect was discovered by the Ancient Greeks (with the help of ancient Babylonian observations). In about 2000 BC, when the science of astronomy originated in ancient Egypt and Babylonia, the vernal equinox lay in the constellation Aries. Indeed, the vernal equinox is still sometimes called the *first point of Aries* in astronomical texts. About 90 BC, the vernal equinox moved into the constellation Pisces, where it still remains. The equinox will move into the constellation Aquarius (marking the beginning of the much heralded "Age of Aquarius") in about 2600 AD. Incidentally, the position of the vernal equinox in the sky is of great significance in astronomy, since it is used as the zero of celestial longitude (much as Greenwich is used as the zero of terrestrial longitude).

Equations (8.117) and (8.118) indicate that the small inclination of the lunar orbit to the ecliptic plane, combined with the precession of the lunar ascending node, causes the Earth's axis of rotation to wobble sightly. This wobble is known as *nutation* (from the Latin *nutare*, to nod), and is superimposed on the aforementioned precession. In the absence of precession, nutation would cause the north celestial pole to periodically trace out a small ellipse on the sky, the sense of rotation being *counter-clockwise*. The nutation period is 18.6 years: *i.e.*, the same as the precession period of the lunar ascending node. The nutation amplitudes in the polar and azimuthal angles θ and ϕ are

$$\delta\theta = \frac{3}{2}\,\frac{\epsilon\,\iota_m\,\mu_m\,T_n(\mathrm{yr})}{T_s(\mathrm{day})\,[T_m(\mathrm{yr})]^2}\,\cos\theta_0, \tag{8.124}$$

$$\delta\phi = \frac{3}{2}\,\frac{\epsilon\,\iota_m\,\mu_m\,T_n(\mathrm{yr})}{T_s(\mathrm{day})\,[T_m(\mathrm{yr})]^2}\,\frac{\cos(2\,\theta_0)}{\sin\theta_0}, \tag{8.125}$$

respectively, where $T_n(\mathrm{yr}) = n_s/\Omega_n = 18.6$. Given that $\epsilon = 0.00335$, $\theta_0 = 23.44°$, $\iota_m = 5.16°$, $T_s(\mathrm{day}) = 365.24$, $T_m(\mathrm{yr}) = 0.081$, and $\mu_m = 0.0123$, we obtain

$$\delta\theta = 8.2'', \tag{8.126}$$

$$\delta\phi = 15.3''. \tag{8.127}$$

The observed nutation amplitudes are $9.2''$ and $17.2''$, respectively. Hence, our estimates are quite close to the mark. Any inaccuracy is mainly due to the fact that we have neglected to take into account the small eccentricities of the Earth's orbit around the Sun, and the Moon's orbit around the Earth. The nutation of the Earth was discovered in 1728 by the English astronomer James Bradley, and was explained theoretically about 20 years later by d'Alembert and Euler. Nutation is important because the corresponding gyration of the Earth's rotation axis appears to be transferred to celestial objects when they are viewed using terrestrial telescopes. This effect causes the celestial longitudes and latitudes of heavenly objects to oscillate sinusoidally by up to $20''$ (*i.e.*, about the maximum apparent angular size of Saturn) with a period of 18.6 years. It is necessary to correct for this oscillation in order to accurately guide terrestrial telescopes to particular objects.

8.11 Exercises

8.1. A rigid body having an axis of symmetry rotates freely about a fixed point under no torques. If α is the angle between the symmetry axis and the instantaneous axis of rotation, show that the angle between the axis of rotation and the invariable line (the **L** vector) is

$$\tan^{-1}\left[\frac{(I_\parallel - I_\perp)\,\tan\alpha}{I_\parallel + I_\perp\,\tan^2\alpha}\right]$$

where $I_\parallel$ (the moment of inertia about the symmetry axis) is greater than $I_\perp$ (the moment of inertia about an axis normal to the symmetry axis).

8.2. Since the greatest value of $I_\parallel/I_\perp$ is 2 (symmetrical lamina) show from the previous result that the angle between the angular velocity and angular momentum vectors cannot exceed $\tan^{-1}(1/\sqrt{8}) \simeq 19.5^\circ$. Find the corresponding value of α.

8.3. A thin uniform rod of length l and mass m is constrained to rotate with constant angular velocity ω about an axis passing through the center of the rod, and making an angle α with the rod. Show that the angular momentum about the center of the rod is perpendicular to the rod, and is of magnitude $(m\,l^2\,\omega/12)\,\sin\alpha$. Show that the torque is perpendicular to both the rod and the angular momentum vector, and is of magnitude $(m\,l^2\omega^2/12)\,\sin\alpha\,\cos\alpha$.

8.4. A thin uniform disk of radius a and mass m is constrained to rotate with constant angular velocity ω about an axis passing through its center, and making an angle α with the normal to the disk. Find the angular momentum about the center of the disk, as well as the torque acting on the disk.

8.5. Consider an artificial satellite in a circular orbit of radius L about the Earth. Suppose that the normal to the plane of the orbit subtends an angle θ with the Earth's axis of rotation. By approximating the orbiting satellite as a uniform ring, demonstrate that the Earth's oblateness causes the plane of the satellite's orbit to precess about the Earth's rotational axis at the rate

$$\frac{\dot{\phi}}{\omega} \simeq -\frac{1}{2}\,\epsilon\left(\frac{R}{L}\right)^2\cos\theta.$$

Here, ω is the satellite's orbital angular velocity, $\epsilon = 0.00335$ the Earth's ellipticity, and R the Earth's radius. Note that the Earth's axial moment of inertial is $I_{\parallel} \simeq (1/3)\, M\, R^2$, where M is the mass of the Earth.

8.6. A *sun-synchronous* satellite is one which always passes over a given point on the Earth at the same local solar time. This is achieved by fixing the precession rate of the satellite's orbital plane such that it matches the rate at which the Sun appears to move against the background of the stars. What orbital altitude above the surface of the Earth would such a satellite need to have in order to fly over all latitudes between 50° N and 50° S? Is the direction of the satellite orbit in the same sense as the Earth's rotation (prograde), or the opposite sense (retrograde)?

9 Three-Body Problem

9.1 Introduction

We saw earlier, in Section 2.9, that an isolated dynamical system consisting of two freely moving point masses exerting forces on one another—which is usually referred to as a *two-body problem*—can always be converted into an equivalent one-body problem. In particular, this implies that we can *exactly solve* a dynamical system containing *two* gravitationally interacting point masses, since the equivalent one-body problem is exactly soluble—see Sections 2.9 and 4.11. What about a system containing *three* gravitationally interacting point masses? Despite hundreds of years of research, no exact solution of this famous problem—which is generally known as the *three-body problem*—has ever been found. It is, however, possible to make some progress by severely restricting the problem's scope.

9.2 Circular Restricted Three-Body Problem

Consider an isolated dynamical system consisting of three gravitationally interacting point masses, m_1, m_2, and m_3. Suppose, however, that the third mass, m_3, is so much smaller than the other two that it has a negligible effect on their motion. Suppose, further, that the first two masses, m_1 and m_2, execute circular orbits about their common center of mass. In the following, we shall investigate this simplified problem, which is generally known as the *circular restricted three-body problem*.

Let us define a Cartesian coordinate system (ξ, η, ζ) in an inertial reference frame whose origin coincides with the center of mass, C, of the two orbiting masses. Furthermore, let the orbital plane of these masses coincide with the ξ-η plane, and let them both lie on the ξ-axis at time $t = 0$—see Figure 9.1. Suppose that R is the constant distance between the two orbiting masses, r_1 the constant distance between mass m_1 and the origin, and r_2 the constant distance between mass m_2 and the origin. Moreover, let ω be the constant orbital angular velocity. It follows, from Section 4.11, that

$$\omega^2 = \frac{G\,M}{R^3}, \tag{9.1}$$

$$\frac{r_1}{r_2} = \frac{m_2}{m_1}, \tag{9.2}$$

where $M = m_1 + m_2$.

It is convenient to choose our unit of length such that $R = 1$, and our unit of mass such that $G\,M = 1$. It follows, from Equation (9.1), that $\omega = 1$. However, we shall continue to retain ω in our equations, for the sake of clarity. Let $\mu_1 = G\,m_1$, and $\mu_2 = G\,m_2 = 1-\mu_1$. It is easily demonstrated that $r_1 = \mu_2$, and $r_2 = 1 - r_1 = \mu_1$. Hence, the two orbiting masses,

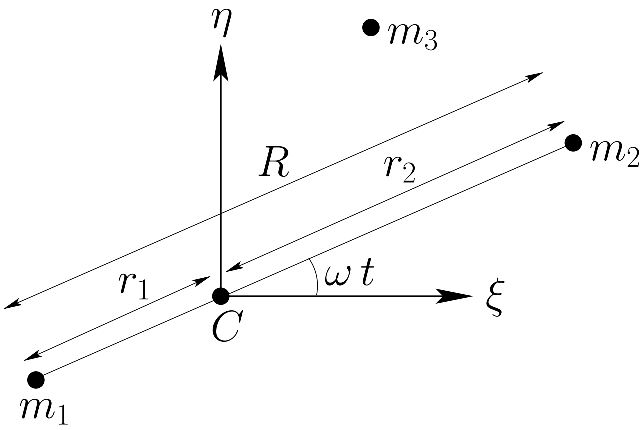


Figure 9.1: *The circular restricted three-body problem.*

m_1 and m_2, have position vectors $\mathbf{r}_1 = (\xi_1, \eta_1, 0)$ and $\mathbf{r}_2 = (\xi_2, \eta_2, 0)$, respectively, where (see Figure 9.1)

$$\mathbf{r}_1 = \mu_2\,(-\cos\omega t,\, -\sin\omega t,\, 0), \tag{9.3}$$

$$\mathbf{r}_2 = \mu_1\,(\cos\omega t,\, \sin\omega t,\, 0). \tag{9.4}$$

Let the third mass have position vector $\mathbf{r} = (\xi, \eta, \zeta)$. The Cartesian components of the equation of motion of this mass are thus

$$\ddot{\xi} = -\mu_1\,\frac{(\xi-\xi_1)}{\rho_1^3} - \mu_2\,\frac{(\xi-\xi_2)}{\rho_2^3}, \tag{9.5}$$

$$\ddot{\eta} = -\mu_1\,\frac{(\eta-\eta_1)}{\rho_1^3} - \mu_2\,\frac{(\eta-\eta_2)}{\rho_2^3}, \tag{9.6}$$

$$\ddot{\zeta} = -\mu_1\,\frac{\zeta}{\rho_1^3} - \mu_2\,\frac{\zeta}{\rho_2^3}, \tag{9.7}$$

where

$$\rho_1^2 = (\xi-\xi_1)^2 + (\eta-\eta_1)^2 + \zeta^2, \tag{9.8}$$

$$\rho_2^2 = (\xi-\xi_2)^2 + (\eta-\eta_2)^2 + \zeta^2. \tag{9.9}$$

9.3 Jacobi Integral

Consider the function

$$C = 2\left(\frac{\mu_1}{\rho_1} + \frac{\mu_2}{\rho_2}\right) + 2\,\omega\,(\xi\,\dot{\eta} - \eta\,\dot{\xi}) - \dot{\xi}^2 - \dot{\eta}^2 - \dot{\zeta}^2. \tag{9.10}$$

The time derivative of this function is written

$$\dot{C} = -\frac{2\,\mu_1\,\dot{\rho}_1}{\rho_1^2} - \frac{2\,\mu_2\,\dot{\rho}_2}{\rho_2^2} + 2\,\omega\,(\xi\,\ddot{\eta} - \eta\,\ddot{\xi}) - 2\,\dot{\xi}\,\ddot{\xi} - 2\,\dot{\eta}\,\ddot{\eta} - 2\,\dot{\zeta}\,\ddot{\zeta}. \tag{9.11}$$

Moreover, it follows, from Equations (9.3)–(9.4) and (9.8)–(9.9), that

$$\rho_1\,\dot{\rho}_1 = -(\xi_1\,\dot{\xi} + \eta_1\,\dot{\eta}) + \omega\,(\xi\,\eta_1 - \eta\,\xi_1) + \xi\,\dot{\xi} + \eta\,\dot{\eta} + \zeta\,\dot{\zeta}, \tag{9.12}$$

$$\rho_2\,\dot{\rho}_2 = -(\xi_2\,\dot{\xi} + \eta_2\,\dot{\eta}) + \omega\,(\xi\,\eta_2 - \eta\,\xi_2) + \xi\,\dot{\xi} + \eta\,\dot{\eta} + \zeta\,\dot{\zeta}. \tag{9.13}$$

Combining Equations (9.5)–(9.7) with the above three expressions, we obtain (after considerable algebra)

$$\dot{C} = 0. \tag{9.14}$$

In other words, the function C—which is usually referred to as the *Jacobi integral*—is a *constant of the motion*.

Now, we can rearrange Equation (9.10) to give

$$\mathcal{E} \equiv \frac{1}{2}\,(\dot{\xi}^2 + \dot{\eta}^2 + \dot{\zeta}^2) - \left(\frac{\mu_1}{\rho_1} + \frac{\mu_2}{\rho_2}\right) = \boldsymbol{\omega}\cdot\mathbf{h} - \frac{C}{2}, \tag{9.15}$$

where $\mathcal{E}$ is the energy (per unit mass) of mass m_3, $\mathbf{h} = \mathbf{r}\times\dot{\mathbf{r}}$ its angular momentum (per unit mass), and $\boldsymbol{\omega} = (0,\,0,\,\omega)$ the orbital angular velocity of the other two masses. Note, however, that $\mathbf{h}$ is *not* a constant of the motion. Hence, $\mathcal{E}$ is not a constant of the motion either. In fact, the Jacobi integral is the *only* constant of the motion in the circular restricted three-body problem. Incidentally, the energy of mass m_3 is not a conserved quantity because the other two masses in the system are *moving*.

9.4 Tisserand Criterion

Consider a dynamical system consisting of three gravitationally interacting point masses, m_1, m_2, and m_3. Let mass m_1 represent the Sun, mass m_2 the planet Jupiter, and mass m_3 a comet. Since the mass of a comet is very much less than that of the Sun or Jupiter, and the Sun and Jupiter are in (almost) circular orbits about their common center of mass, the dynamical system in question satisfies all of the necessary criteria to be considered an example of a restricted three-body problem.

Now, the mass of the Sun is much greater than that of Jupiter. It follows that the gravitational effect of Jupiter on the cometary orbit is *negligible* unless the comet makes a very *close approach* to Jupiter. Hence, as described in Chapter 4, before and after such an approach, the comet executes a standard elliptical orbit about the Sun with fixed orbital parameters: *i.e.*, fixed major radius, eccentricity, and inclination to the ecliptic plane. However, in general, the orbital parameters before and after the close approach will *not* be the same as one another. Let us investigate further.

Now, since $m_1 \gg m_2$, we have $\mu_1 = G\, m_1 \simeq G\,(m_1 + m_2) = 1$, and $\rho_1 \simeq r$. Hence, according to Equations (4.35) and (4.44), the (approximately) conserved energy (per unit mass) of the comet before and after its close approach to Jupiter is

$$\mathcal{E} \equiv \frac{1}{2}\left(\dot{\xi}^2 + \dot{\eta}^2 + \dot{\zeta}^2\right) - \frac{1}{r} = -\frac{1}{2\,a}. \tag{9.16}$$

Note that the comet's orbital energy is entirely determined by its major radius, a. (Incidentally, we are working in units such that the major radius of Jupiter's orbit is unity.) Furthermore, the (approximately) conserved angular momentum (per unit mass) of the comet before and after its approach to Jupiter is written $\mathbf{h}$, where $\mathbf{h}$ is directed *normal* to the comet's orbital plane, and, from Equations (B.7) and (4.31),

$$h^2 = a\,(1 - e^2). \tag{9.17}$$

Here, e is the comet's orbital eccentricity. It follows that

$$\boldsymbol{\omega} \cdot \mathbf{h} = \omega\, h\, \cos I = \sqrt{a\,(1 - e^2)}\, \cos I, \tag{9.18}$$

since $\omega = 1$ in our adopted system of units. Here, I is the angle of inclination of the normal to the comet's orbital plane to that of Jupiter's orbital plane.

Let a, e, and I be the major radius, eccentricity, and inclination angle of the cometary orbit before the close encounter with Jupiter, and let a', e', and I' be the corresponding parameters after the encounter. It follows from Equations (9.15), (9.16), and (9.18), and the fact that C is conserved during the encounter, whereas $\mathcal{E}$ and h are not, that

$$\frac{1}{2\,a} + \sqrt{a\,(1 - e^2)}\, \cos I = \frac{1}{2\,a'} + \sqrt{a'\,(1 - e'^2)}\, \cos I'. \tag{9.19}$$

This result is known as the *Tisserand criterion*, after the French astronomer Felix Tisserand (1845-1896) who discovered it, and restricts the possible changes in the orbital parameters of a comet due to a close encounter with Jupiter (or any other massive planet).

The Tisserand criterion is very useful. For instance, whenever a new comet is discovered, astronomers immediately calculate its *Tisserand parameter*,

$$T_J = \frac{1}{a} + 2\,\sqrt{a\,(1 - e^2)}\, \cos I. \tag{9.20}$$

If this parameter has the same value as that of a previously observed comet then it is quite likely that the new comet is, in fact, the same comet, but that its orbital parameters have changed since it was last observed, due to a close encounter with Jupiter. Incidentally, the subscript J in the above formula is to remind us that we are dealing with the Tisserand parameter for close encounters with *Jupiter*. (The parameter is, thus, evaluated in a system of units in which the major radius of Jupiter's orbit is unity). Obviously, it is also possible to calculate Tisserand parameters for close encounters with other planets.

The Tisserand criterion is also applicable to so-called *gravity assists*, in which a spacecraft gains energy due to a close encounter with a moving planet. Such assists are often

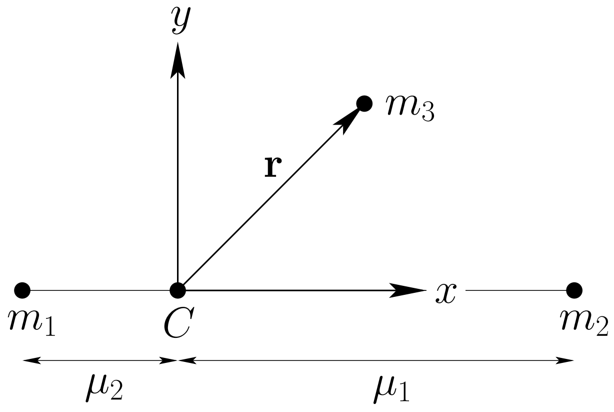


Figure 9.2: *The co-rotating frame.*

employed in missions to the outer planets to reduce the amount of fuel which the space-craft must carry in order to reach its destination. In fact, it is clear, from Equations (9.16) and (9.19), that a space-craft can make use of a close encounter with a moving planet to increase (or decrease) its orbital major radius a, and, hence, to increase (or decrease) its total orbital energy.

9.5 Co-Rotating Frame

Let us transform to a non-inertial frame of reference rotating with angular velocity ω about an axis normal to the orbital plane of masses m_1 and m_2, and passing through their center of mass. It follows that masses m_1 and m_2 appear *stationary* in this new reference frame. Let us define a Cartesian coordinate system (x, y, z) in the rotating frame of reference which is such that masses m_1 and m_2 always lie on the x-axis, and the z-axis is parallel to the previously defined ζ-axis. It follows that masses m_1 and m_2 have the fixed position vectors $\mathbf{r}_1 = \mu_2\,(-1, 0, 0)$ and $\mathbf{r}_2 = \mu_1\,(1, 0, 0)$ in our new coordinate system. Finally, let the position vector of mass m_3 be $\mathbf{r} = (x, y, z)$—see Figure 9.2.

According to Section 6.2, the equation of motion of mass m_3 in the rotating reference frame takes the form

$$\ddot{\mathbf{r}} + 2\,\boldsymbol{\omega} \times \dot{\mathbf{r}} = -\mu_1\,\frac{(\mathbf{r} - \mathbf{r}_1)}{\rho_1^3} - \mu_2\,\frac{(\mathbf{r} - \mathbf{r}_2)}{\rho_2^3} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}), \tag{9.21}$$

where $\boldsymbol{\omega} = (0, 0, \omega)$, and

$$\rho_1^2 = (x + \mu_2)^2 + y^2 + z^2, \tag{9.22}$$

$$\rho_2^2 = (x - \mu_1)^2 + y^2 + z^2. \tag{9.23}$$

Here, the second term on the left-hand side of Equation (9.21) is the *Coriolis* acceleration, whereas the final term on the right-hand side is the *centrifugal* acceleration. The

components of Equation (9.21) reduce to

$$\ddot{x} - 2\,\omega\,\dot{y} = -\frac{\mu_1\,(x+\mu_2)}{\rho_1^3} - \frac{\mu_2\,(x-\mu_1)}{\rho_2^3} + \omega^2\,x, \tag{9.24}$$

$$\ddot{y} + 2\,\omega\,\dot{x} = -\frac{\mu_1\,y}{\rho_1^3} - \frac{\mu_2\,y}{\rho_2^3} + \omega^2\,y, \tag{9.25}$$

$$\ddot{z} = -\frac{\mu_1\,z}{\rho_1^3} - \frac{\mu_2\,z}{\rho_2^3}, \tag{9.26}$$

which yield

$$\ddot{x} - 2\,\omega\,\dot{y} = -\frac{\partial U}{\partial x}, \tag{9.27}$$

$$\ddot{y} + 2\,\omega\,\dot{x} = -\frac{\partial U}{\partial y}, \tag{9.28}$$

$$\ddot{z} = -\frac{\partial U}{\partial z}, \tag{9.29}$$

where

$$U = -\frac{\mu_1}{\rho_1} - \frac{\mu_2}{\rho_2} - \frac{\omega^2}{2}\,(x^2+y^2) \tag{9.30}$$

is the sum of the gravitational and centrifugal potentials.

Now, it follows from Equations (9.27)–(9.29) that

$$\ddot{x}\,\dot{x} - 2\,\omega\,\dot{x}\,\dot{y} = -\dot{x}\,\frac{\partial U}{\partial x}, \tag{9.31}$$

$$\ddot{y}\,\dot{y} + 2\,\omega\,\dot{x}\,\dot{y} = -\dot{y}\,\frac{\partial U}{\partial y}, \tag{9.32}$$

$$\ddot{z}\,\dot{z} = -\dot{z}\,\frac{\partial U}{\partial z}. \tag{9.33}$$

Summing the above three equations, we obtain

$$\frac{d}{dt}\left[\frac{1}{2}\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)+U\right] = 0. \tag{9.34}$$

In other words,

$$C = -2\,U - v^2 \tag{9.35}$$

is a *constant of the motion*, where $v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$. In fact, C is the *Jacobi integral* introduced in Section 9.3 [it is easily demonstrated that Equations (9.10) and (9.35) are identical]. Note, finally, that the mass m_3 is restricted to regions in which

$$-2\,U \geq C, \tag{9.36}$$

since v^2 is a positive definite quantity.

9.6 Lagrange Points

Let us search for possible *equilibrium points* of the mass m_3 in the rotating reference frame. Such points are termed *Lagrange points*. Thus, in the rotating frame, the mass m_3 would remain at rest if placed at one of the Lagrange points. It is, thus, clear that these points are *fixed* in the rotating frame. Conversely, in the inertial reference frame, the Lagrange points *rotate* about the center of mass with angular velocity ω, and the mass m_3 would consequently also rotate about the center of mass with angular velocity ω if placed at one of these points (with the appropriate velocity). In the following, we shall assume, without loss of generality, that $m_1 \geq m_2$.

The Lagrange points satisfy $\dot{\mathbf{r}} = \ddot{\mathbf{r}} = \mathbf{0}$ in the rotating frame. It thus follows, from Equations (9.27)–(9.29), that the Lagrange points are the solutions of

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial y} = \frac{\partial U}{\partial z} = 0. \tag{9.37}$$

Now, it is easily seen that

$$\frac{\partial U}{\partial z} = \left(\frac{\mu_1}{\rho_1^3} + \frac{\mu_2}{\rho_2^3}\right) z. \tag{9.38}$$

Since the term in brackets is positive definite, we conclude that the only solution to the above equation is $z = 0$. Hence, all of the Lagrange points lie in the x-y plane.

If $z = 0$ then it is readily demonstrated that

$$\mu_1\,\rho_1^2 + \mu_2\,\rho_2^2 = x^2 + y^2 + \mu_1\,\mu_2, \tag{9.39}$$

where use has been made of the fact that $\mu_1 + \mu_2 = 1$. Hence, Equation (9.30) can also be written

$$U = -\mu_1\left(\frac{1}{\rho_1} + \frac{\rho_1^2}{2}\right) - \mu_2\left(\frac{1}{\rho_2} + \frac{\rho_2^2}{2}\right) + \frac{\mu_1\,\mu_2}{2}. \tag{9.40}$$

The Lagrange points thus satisfy

$$\frac{\partial U}{\partial x} = \frac{\partial U}{\partial \rho_1}\frac{\partial \rho_1}{\partial x} + \frac{\partial U}{\partial \rho_2}\frac{\partial \rho_2}{\partial x} = 0, \tag{9.41}$$

$$\frac{\partial U}{\partial y} = \frac{\partial U}{\partial \rho_1}\frac{\partial \rho_1}{\partial y} + \frac{\partial U}{\partial \rho_2}\frac{\partial \rho_2}{\partial y} = 0, \tag{9.42}$$

which reduce to

$$\mu_1\left(\frac{1-\rho_1^3}{\rho_1^2}\right)\left(\frac{x+\mu_2}{\rho_1}\right) + \mu_2\left(\frac{1-\rho_2^3}{\rho_2^2}\right)\left(\frac{x-\mu_1}{\rho_2}\right) = 0, \tag{9.43}$$

$$\mu_1\left(\frac{1-\rho_1^3}{\rho_1^2}\right)\left(\frac{y}{\rho_1}\right) + \mu_2\left(\frac{1-\rho_2^3}{\rho_2^2}\right)\left(\frac{y}{\rho_2}\right) = 0. \tag{9.44}$$

Now, one obvious solution of Equation (9.44) is $y = 0$, corresponding to a Lagrange point which lies on the x-axis. It turns out that there are *three* such points. L_1 lies between

masses m_1 and m_2, L_2 lies to the right of mass m_2, and L_3 lies to the left of mass m_1—see Figure 9.2. At the L_1 point, we have $x = -\mu_2 + \rho_1 = \mu_1 - \rho_2$ and $\rho_1 = 1 - \rho_2$. Hence, from Equation (9.43),

$$\frac{\mu_2}{3\,\mu_1} = \frac{\rho_2^3\,(1 - \rho_2 + \rho_2^2/3)}{(1 + \rho_2 + \rho_2^2)\,(1 - \rho_2)^3}. \tag{9.45}$$

Assuming that $\rho_2 \ll 1$, we can find an approximate solution of Equation (9.45) by expanding in powers of ρ_2:

$$\alpha = \rho_2 + \frac{\rho_2^2}{3} + \frac{\rho_2^3}{3} + \frac{51\,\rho_2^4}{81} + \mathcal{O}(\rho_2^5). \tag{9.46}$$

This equation can be inverted to give

$$\rho_2 = \alpha - \frac{\alpha^2}{3} - \frac{\alpha^3}{9} - \frac{23\,\alpha^4}{81} + \mathcal{O}(\alpha^5), \tag{9.47}$$

where

$$\alpha = \left(\frac{\mu_2}{3\,\mu_1}\right)^{1/3}. \tag{9.48}$$

is assumed to be a small parameter.

At the L_2 point, we have $x = -\mu_2 + \rho_1 = \mu_1 + \rho_2$ and $\rho_1 = 1 + \rho_2$. Hence, from Equation (9.43),

$$\frac{\mu_2}{3\,\mu_1} = \frac{\rho_2^3\,(1 + \rho_2 + \rho_2^2/3)}{(1 + \rho_2)^2\,(1 - \rho_2^3)}. \tag{9.49}$$

Again, expanding in powers of ρ_2, we obtain

$$\alpha = \rho_2 - \frac{\rho_2^2}{3} + \frac{\rho_2^3}{3} + \frac{\rho_2^4}{81} + \mathcal{O}(\rho_2^5), \tag{9.50}$$

$$\rho_2 = \alpha + \frac{\alpha^2}{3} - \frac{\alpha^3}{9} - \frac{31\,\alpha^4}{81} + \mathcal{O}(\alpha^5). \tag{9.51}$$

Finally, at the L_3 point, we have $x = -\mu_2 - \rho_1 = \mu_1 - \rho_2$ and $\rho_2 = 1 + \rho_1$. Hence, from Equation (9.43),

$$\frac{\mu_2}{\mu_1} = \frac{(1 - \rho_1^3)\,(1 + \rho_1)^2}{\rho_1^3\,(\rho_1^2 + 3\,\rho_1 + 3)}. \tag{9.52}$$

Let $\rho_1 = 1 - \beta$. Expanding in powers of β, we obtain

$$\frac{\mu_2}{\mu_1} = \frac{12\,\beta}{7} + \frac{144\,\beta^2}{49} + \frac{1567\,\beta^3}{343} + \mathcal{O}(\beta^4), \tag{9.53}$$

$$\beta = \frac{7}{12}\left(\frac{\mu_2}{\mu_1}\right) - \frac{7}{12}\left(\frac{\mu_2}{\mu_1}\right)^2 + \frac{13223}{20736}\left(\frac{\mu_2}{\mu_1}\right)^3 + \mathcal{O}\left(\frac{\mu_2}{\mu_1}\right)^4, \tag{9.54}$$

where μ_2/μ_1 is assumed to be a small parameter.

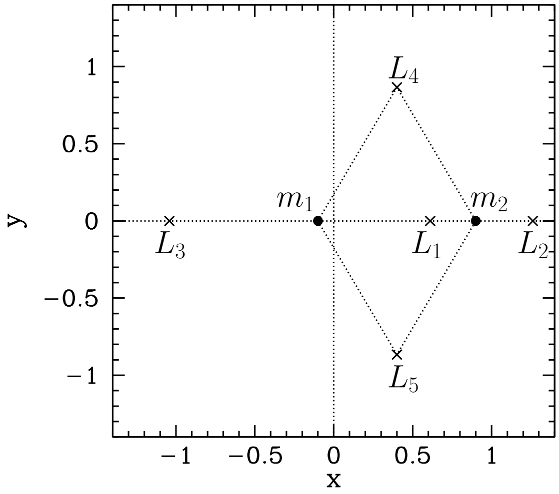


Figure 9.3: *The masses* m_1 *and* m_2*, and the five Lagrange points,* L_1 *to* L_5*, calculated for* $\mu_2 = 0.1$.

Let us now search for Lagrange points which *do not lie* on the x-axis. One obvious solution of Equations (9.41) and (9.42) is

$$\frac{\partial U}{\partial \rho_1} = \frac{\partial U}{\partial \rho_2} = 0, \tag{9.55}$$

giving, from Equation (9.40),

$$\rho_1 = \rho_2 = 1, \tag{9.56}$$

or

$$(x + \mu_2)^2 + y^2 = (x - 1 + \mu_2)^2 + y^2 = 1, \tag{9.57}$$

since $\mu_1 = 1 - \mu_2$. The two solutions of the above equation are

$$x = \frac{1}{2} - \mu_2, \tag{9.58}$$

$$y = \pm\frac{\sqrt{3}}{2}, \tag{9.59}$$

and specify the positions of the Lagrange points designated L_4 and L_5. Note that point L_4 and the masses m_1 and m_2 lie at the apexes of an *equilateral triangle*. The same is true for point L_5. We have now found all of the possible Lagrange points.

Figure 9.3 shows the positions of the two masses, m_1 and m_2, and the five Lagrange points, L_1 to L_5, calculated for the case where $\mu_2 = 0.1$.

9.7 Zero-Velocity Surfaces

Consider the surface

$$V(x, y, z) = C, \tag{9.60}$$

where

$$V = -2\,U = \frac{2\,\mu_1}{\rho_1} + \frac{2\,\mu_2}{\rho_2} + x^2 + y^2. \tag{9.61}$$

Note that $V \geq 0$. It follows, from Equation (9.35), that if the mass m_3 has the Jacobi integral C, and lies on the surface specified in Equation (9.60), then it must have *zero* velocity. Hence, such a surface is termed a *zero-velocity surface*. The zero-velocity surfaces are important because they form the boundary of regions from which the mass m_3 is dynamically *excluded*: *i.e.*, regions in which $V < C$. Generally speaking, the regions from which m_3 is excluded grow in area as C increases, and *vice versa*.

Let C_i be the value of V at the L_i Lagrange point, for $i = 1, 5$. When $\mu_2 \ll 1$, it is easily demonstrated that

$$C_1 \simeq 3 + 3^{4/3}\,\mu_2^{2/3} - 10\,\mu_2/3, \tag{9.62}$$

$$C_2 \simeq 3 + 3^{4/3}\,\mu_2^{2/3} - 14\,\mu_2/3, \tag{9.63}$$

$$C_3 \simeq 3 + \mu_2, \tag{9.64}$$

$$C_4 \simeq 3 - \mu_2, \tag{9.65}$$

$$C_4 \simeq 3 - \mu_2. \tag{9.66}$$

Note that $C_1 > C_2 > C_3 > C_4 = C_5$.

Figures 9.4–9.8 show the intersection of the zero-velocity surface $V = C$ with the x-y plane for various different values of C, and illustrate how the region from which m_3 is dynamically excluded—which we shall term the *excluded region*—evolves as the value of C is varied. Of course, any point not in the excluded region is in the so-called *allowed region*. For $C > C_1$, the allowed region consists of two separate oval regions centered on m_1 and m_2, respectively, plus an outer region which lies beyond a large circle centered on the origin. All three allowed regions are separated from one another by an excluded region—see Figure 9.4. When $C = C_1$, the two inner allowed regions merge at the L_1 point—see Figure 9.5. When $C = C_2$, the inner and outer allowed regions merge at the L_2 point, forming a horseshoe-like excluded region—see Figure 9.6. When $C = C_3$, the excluded region splits in two at the L_3 point—see Figure 9.7. For $C_4 < C < C_3$, the two excluded regions are localized about the L_4 and L_5 points—see Figure 9.8. Finally, for $C < C_4$, there is no excluded region.

Figure 9.9 shows the zero-velocity surfaces and Lagrange points calculated for the case $\mu_2 = 0.01$. It can be seen that, at very small values of μ_2, the L_1 and L_2 Lagrange points are almost *equidistant* from mass m_2. Furthermore, mass m_2, and the L_3, L_4, and L_5 Lagrange points all lie approximately on a *unit circle*, centered on mass m_1. It follows that, when μ_2 is small, the Lagrange points L_3, L_4 and L_5 all share the orbit of mass m_2 about m_1 (in the

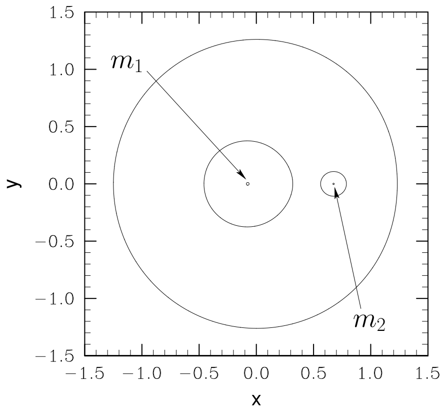


Figure 9.4: *The zero-velocity surface* $V = C$*, where* $C > C_1$*, calculated for* $\mu_2 = 0.1$*. The mass* m_3 *is excluded from the region lying between the two inner curves and the outer curve.*

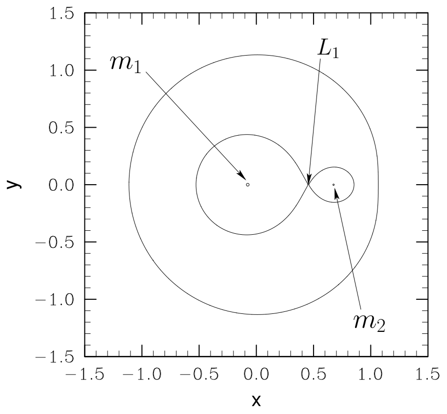


Figure 9.5: *The zero-velocity surface* $V = C$*, where* $C = C_1$*, calculated for* $\mu_2 = 0.1$*. The mass* m_3 *is excluded from the region lying between the inner and outer curves.*

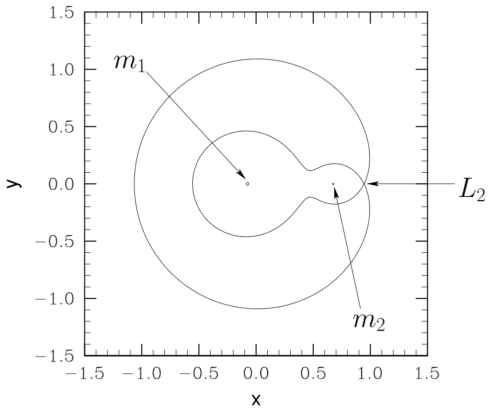


Figure 9.6: *The zero-velocity surface* $V = C$*, where* $C = C_2$*, calculated for* $\mu_2 = 0.1$*. The mass* m_3 *is excluded from the region lying between the inner and outer curve.*

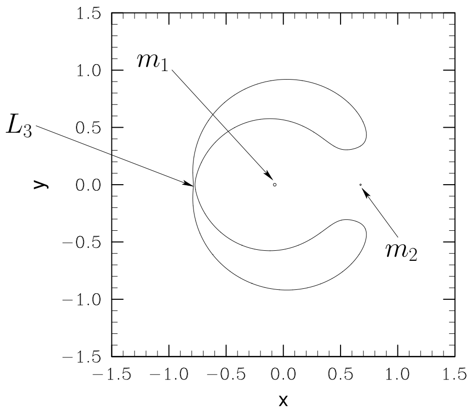


Figure 9.7: *The zero-velocity surface* $V = C$*, where* $C = C_3$*, calculated for* $\mu_2 = 0.1$*. The mass* m_3 *is excluded from the regions lying inside the curve.*

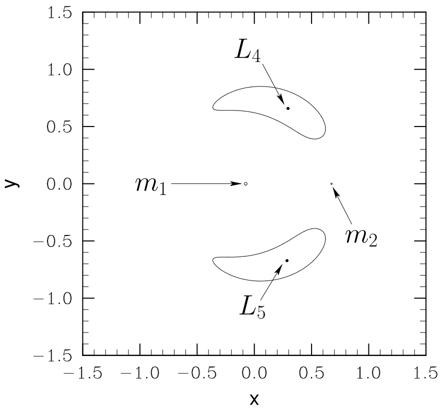


Figure 9.8: *The zero-velocity surface* $V = C$*, where* $C_4 < C < C_3$*, calculated for* $\mu_2 = 0.1$*. The mass* m_3 *is excluded from the regions lying inside the two curves.*

inertial frame) with C_3 being directly opposite m_2, L_4 (by convention) 60° ahead of m_2, and L_5 60° behind.

9.8 Stability of Lagrange Points

We have seen that the five Lagrange points, L_1 to L_5, are the equilibrium points of mass m_3 in the co-rotating frame. Let us now determine whether or not these equilibrium points are *stable* to small displacements.

Now, the equations of motion of mass m_3 in the co-rotating frame are specified in Equations (9.27)–(9.29). Note that the motion in the x-y plane is complicated by presence of the Coriolis acceleration. However, the motion parallel to the z-axis simply corresponds to motion in the potential U. Hence, the condition for the stability of the Lagrange points (which all lie at $z = 0$) to small displacements parallel to the z-axis is simply (see Section 2.7)

$$\left(\frac{\partial^2 U}{\partial z^2}\right)_{z=0} = \frac{\mu_1}{\rho_1^3} + \frac{\mu_2}{\rho_2^3} > 0. \tag{9.67}$$

This condition is satisfied everywhere in the x-y plane. Hence, the Lagrange points are all *stable* to small displacements parallel to the z-axis. It, thus, remains to investigate their stability to small displacements lying within the x-y plane.

Suppose that a Lagrange point is situated in the x-y plane at coordinates $(x_0, y_0, 0)$. Let us consider small amplitude x-y motion in the vicinity of this point by writing

$$x = x_0 + \delta x, \tag{9.68}$$

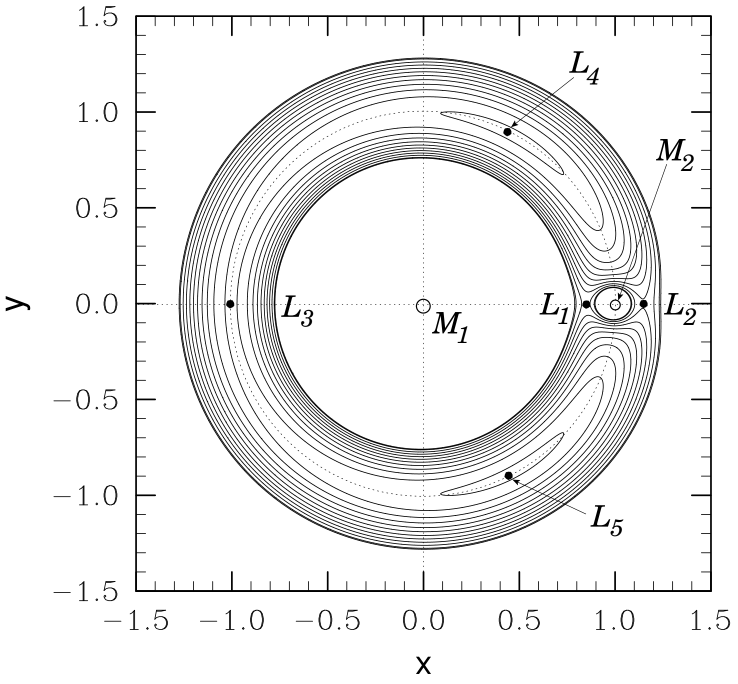


Figure 9.9: *The zero-velocity surfaces and Lagrange points calculated for* $\mu_2 = 0.01$.

$$y = y_0 + \delta y, \tag{9.69}$$

$$z = 0, \tag{9.70}$$

where δx and δy are infinitesimal. Expanding $U(x, y, 0)$ about the Lagrange point as a Taylor series, and retaining terms up to second-order in small quantities, we obtain

$$U = U_0 + U_x\,\delta x + U_y\,\delta y + \frac{1}{2}\,U_{xx}\,(\delta x)^2 + U_{xy}\,\delta x\,\delta y + \frac{1}{2}\,U_{yy}\,(\delta y)^2, \tag{9.71}$$

where $U_0 = U(x_0, y_0, 0)$, $U_x = \partial U(x_0, y_0, 0)/\partial x$, $U_{xx} = \partial^2 U(x_0, y_0, 0)/\partial x^2$, *etc.* However, by definition, $U_x = U_y = 0$ at a Lagrange point, so the expansion simplifies to

$$U = U_0 + \frac{1}{2}\,U_{xx}\,(\delta x)^2 + U_{xy}\,\delta x\,\delta y + \frac{1}{2}\,U_{yy}\,(\delta y)^2. \tag{9.72}$$

Finally, substitution of Equations (9.68)–(9.70), and (9.72) into the equations of x-y motion, (9.27) and (9.28), yields

$$\delta\ddot{x} - 2\,\delta\dot{y} = -U_{xx}\,\delta x - U_{xy}\,\delta y, \tag{9.73}$$

$$\delta\ddot{y} + 2\,\delta\dot{x} = -U_{xy}\,\delta x - U_{yy}\,\delta y, \tag{9.74}$$

since $\omega = 1$.

Let us search for a solution of the above pair of equations of the form $\delta x(t) = \delta x_0\,\exp(\gamma\,t)$ and $\delta y(t) = \delta y_0\,\exp(\gamma\,t)$. We obtain

$$\begin{pmatrix} \gamma^2 + U_{xx} & -2\,\gamma + U_{xy} \\ 2\,\gamma + U_{xy} & \gamma^2 + U_{yy} \end{pmatrix} \begin{pmatrix} \delta x_0 \\ \delta y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{9.75}$$

This equation only has a non-trivial solution if the determinant of the matrix is zero. Hence, we get

$$\gamma^4 + (4 + U_{xx} + U_{yy})\,\gamma^2 + (U_{xx}\,U_{yy} - U_{xy}^2) = 0. \tag{9.76}$$

Now, it is convenient to define

$$A = \frac{\mu_1}{\rho_1^3} + \frac{\mu_2}{\rho_2^3}, \tag{9.77}$$

$$B = 3\left[\frac{\mu_1}{\rho_1^5} + \frac{\mu_2}{\rho_2^5}\right] y^2, \tag{9.78}$$

$$C = 3\left[\frac{\mu_1\,(x + \mu_2)}{\rho_1^5} + \frac{\mu_2\,(x - \mu_1)}{\rho_2^5}\right] y, \tag{9.79}$$

$$D = 3\left[\frac{\mu_1\,(x + \mu_2)^2}{\rho_1^3} + \frac{\mu_2\,(x - \mu_1)^2}{\rho_2^3}\right], \tag{9.80}$$

where all terms are evaluated at the point $(x_0, y_0, 0)$. It thus follows that

$$U_{xx} = A - D - 1, \quad (9.81)$$

$$U_{yy} = A - B - 1, \quad (9.82)$$

$$U_{xy} = -C. \quad (9.83)$$

Consider the co-linear Lagrange points, L_1, L_2, and L_3. These all lie on the x-axis, and are thus characterized by $y = 0$, $\rho_1^2 = (x + \mu_2)^2$, and $\rho_2^2 = (x - \mu_1)^2$. It follows, from the above equations, that $B = C = 0$ and $D = 3\,A$. Hence, $U_{xx} = -1 - 2\,A$, $U_{yy} = A - 1$, and $U_{xy} = 0$. Equation (9.76) thus yields

$$\Gamma^2 + (2 - A)\,\Gamma + (1 - A)\,(1 + 2\,A) = 0, \quad (9.84)$$

where $\Gamma = \gamma^2$. Now, in order for a Lagrange point to be stable to small displacements, all four of the roots, γ, of Equation (9.76) must be *purely imaginary*. This, in turn, implies that the two roots of the above equation,

$$\Gamma = \frac{A - 2 \pm \sqrt{A\,(9\,A - 8)}}{2}, \quad (9.85)$$

must both be *real and negative*. Thus, the stability criterion is

$$\frac{8}{9} \leq A \leq 1. \quad (9.86)$$

Figure 9.10 shows A calculated at the three co-linear Lagrange points as a function of μ_2, for all allowed values of this parameter (*i.e.*, $0 < \mu_2 \leq 0.5$). It can be seen that A is always greater than unity for all three points. Hence, we conclude that the co-linear Lagrange points, L_1, L_2, and L_3, are intrinsically *unstable* equilibrium points in the co-rotating frame.

Let us now consider the triangular Lagrange points, L_4 and L_5. These points are characterized by $\rho_1 = \rho_2 = 1$. It follows that $A = 1$, $B = 9/4$, $C = \pm\sqrt{27/16}\,(1 - 2\,\mu_2)$, and $D = 3/4$. Hence, $U_{xx} = -3/4$, $U_{yy} = -9/4$, and $U_{xy} = \mp\sqrt{27/16}\,(1 - 2\,\mu_2)$, where the upper/lower signs corresponds to L_4 and L_5, respectively. Equation (9.76) thus yields

$$\Gamma^2 + \Gamma + \frac{27}{4}\,\mu_2\,(1 - \mu_2) = 0 \quad (9.87)$$

for both points, where $\Gamma = \gamma^2$. As before, the stability criterion is that the two roots of the above equation must both be real and negative. This is the case provided that $1 > 27\,\mu_2\,(1 - \mu_2)$, which yields the stability criterion

$$\mu_2 < \frac{1}{2}\left(1 - \sqrt{\frac{23}{27}}\right) = 0.0385. \quad (9.88)$$

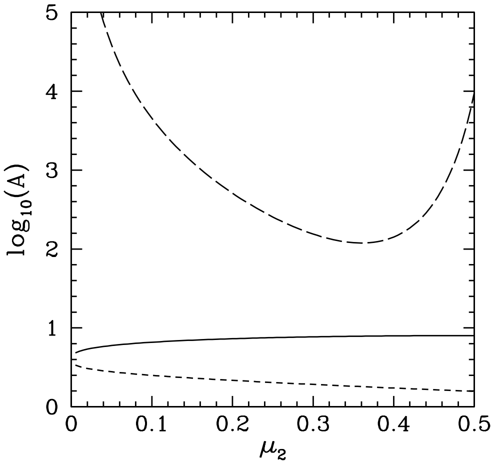


Figure 9.10: *The solid, short-dashed, and long-dashed curves show* A *as a function of* μ_2 *at the* L_1*,* L_2*, and* L_3 *Lagrange points.*

In unnormalized units, this criterion becomes

$$\frac{m_2}{m_1 + m_2} < 0.0385. \tag{9.89}$$

We thus conclude that the L_4 and L_5 Lagrange points are *stable* equilibrium points, in the co-rotating frame, provided that mass m_2 is less than about 4% of mass m_1. If this is the case then mass m_3 can orbit around these points indefinitely. In the inertial frame, the mass will share the orbit of mass m_2 about mass m_1, but will stay approximately 60° *ahead of* mass m_2, if it is orbiting the L_4 point, or 60° *behind*, if it is orbiting the L_5 point—see Figure 9.9. This type of behavior has been observed in the Solar System. For instance, there is a sub-class of asteroids, known as the *Trojan asteroids*, which are trapped in the vicinity of the L_4 and L_5 points of the Sun-Jupiter system [which easily satisfies the stability criterion (9.89)], and consequently share Jupiter's orbit around the Sun, staying approximately 60° *ahead of*, and 60° *behind*, Jupiter, respectively. Furthermore, the L_4 and L_5 points of the Sun-Earth system are occupied by clouds of dust.

10 Lunar Motion

10.1 Historical Background

The motion of the Planets around the Sun is fairly accurately described by *Kepler's laws* (see Chapter 4). Likewise, to a first approximation, the motion of the Moon around the Earth can also be accounted for via these laws. However, unlike the planetary orbits, the deviations of the lunar orbit from a Keplerian ellipse are sufficiently large that they are easily apparent to the naked eye. Indeed, the largest of these deviations, which is generally known as *evection*, was discovered in ancient times by the Alexandrian astronomer Claudius Ptolemy (90–168 AD). Moreover, the next largest deviation, which is called *variation*, was first observed by Tycho Brahe (1546–1601) without the aid of a telescope. Another non-Keplerian feature of the lunar orbit, which is sufficient obvious that it was known to the ancient Greeks, is the fact that the lunar *perigee* (*i.e.*, the point of closest approach to the Earth) *precesses* (*i.e.*, orbits about the Earth in the same direction as the Moon) at such a rate that it completes a full circuit every 8.85 years.[1] The ancient Greeks also noticed that the lunar *ascending node* (*i.e.*, the point at which the Moon passes through the fixed plane of the Earth's orbit around the Sun from south to north) *regresses* (*i.e.*, orbits about the Earth in the opposite direction to the Moon) at such a rate that it completes a full circuit every 18.6 years. Of course, according to standard two-body orbit theory, both the lunar perigee and ascending node should be *stationary* (see Chapter 4).

Newton demonstrated, in Book III of his Principia, that the deviations of the lunar orbit from a Keplerian ellipse are due to the gravitational influence of the Sun, which is sufficiently large that it is not completely negligible compared with the mutual gravitational attraction of the Earth and the Moon. However, Newton was not able to give a complete account of these deviations, due to the complexity of the equations of motion which arise in a system of three mutually gravitating bodies (see Chapter 9). In fact, Clairaut (1713–1765) is generally credited with the first reasonably accurate and complete theoretical explanation of the Moon's orbit. His method of calculation made use of an expansion of the lunar equations of motion in terms of various small parameters. Clairaut, however, initially experienced difficulty in accounting for the precession of the lunar perigee. Indeed, his first calculation overestimated the period of this precession by a factor of about two, leading him to question Newton's inverse-square law of gravitation. Later on, he realized that he could indeed account for the precession, in terms of standard Newtonian dynamics, by continuing his expansion in small parameters to higher order. After Clairaut, the theory of lunar motion was further elaborated in major works by D'Alembert (1717–1783), Euler (1707–1783), Laplace (1749–1827), Damoiseau (1768–1846), Plana (1781–1864), Poisson (1781–1840), Hansen (1795–1874), De Pontécoulant (1795–1824),

[1] Note that this precession rate is about 10^4 times greater than any of the planetary perihelion precession rates discussed in Section 5.4.

J. Herschel (1792–1871), Airy (1801–1892), Delaunay (1816–1872), G.W. Hill (1836–1914), and E.W. Brown (1836–1938). The fact that so many celebrated mathematicians and astronomers devoted so much time and effort to lunar theory is a tribute to its inherent difficulty, as well as its great theoretical and practical interest. Indeed, for a period of about one hundred years (between 1767 and about 1850) the so-called *method of lunar distance* was the principal means used by mariners to determine longitude at sea. This method depends crucially on a precise knowledge of the position of the Moon in the sky as a function of time. Consequently, astronomers and mathematicians during the period in question were spurred to make ever more accurate observations of the Moon's orbit, and to develop lunar theory to greater and greater precision. An important outcome of these activities were various tables of lunar motion (*e.g.*, those of Mayer, Damoiseau, Plana, Hansen, and Brown), most of which were published at public expense.

This chapter contains an introduction to lunar theory in which approximate expressions for evection, variation, the precession of the perigee, and the regression of the ascending node, are derived from the laws of Newtonian mechanics.

10.2 Preliminary Analysis

Let $\mathbf{r}_E$ and $\mathbf{r}_M$ be the position vectors of the Earth and Moon, respectively, in a *non-rotating* reference frame in which the Sun is at rest at the origin. Treating this reference frame as *inertial* (which is an excellent approximation, given that the mass of the Sun is very much greater than that of the Earth or the Moon), the Earth's equation of motion becomes (see Chapter 4)

$$\ddot{\mathbf{r}}_E = -n'^2\, a'^3\, \frac{\mathbf{r}_E}{|\mathbf{r}_E|^3}, \tag{10.1}$$

where $n' = 0.98560025^\circ$ per day and $a' = 149,598,261$ km are the mean angular velocity and major radius, respectively, of the terrestrial orbit about the Sun. Here, $\ddot{} \equiv d^2/dt^2$. On the other hand, the Moon's equation of motion takes the form

$$\ddot{\mathbf{r}}_M = -n^2\, a^3\, \frac{(\mathbf{r}_M - \mathbf{r}_E)}{|\mathbf{r}_M - \mathbf{r}_E|^3} - n'^2\, a'^3\, \frac{\mathbf{r}_M}{|\mathbf{r}_M|^3}, \tag{10.2}$$

where $n = 13.17639646^\circ$ per day and $a = 384,399$ km are the mean angular velocity and major radius, respectively, of the lunar orbit about the Earth. Note that we have retained the acceleration due to the Earth in the lunar equation of motion, (10.2), whilst neglecting the acceleration due to the Moon in the terrestrial equation of motion, (10.1), since the former acceleration is significantly greater (by a factor $M_E/M_M \simeq 81$, where M_E is the mass of the Earth, and M_M the mass of the Moon) than the latter.

Let

$$\mathbf{r} = \mathbf{r}_M - \mathbf{r}_E, \tag{10.3}$$

$$\mathbf{r}' = -\mathbf{r}_E, \tag{10.4}$$

be the position vectors of the Moon and Sun, respectively, relative to the Earth. It follows, from Equations (10.1)–(10.4), that in a non-inertial reference frame, S (say), in which the Earth is at rest at the origin, but the coordinate axes point in *fixed* directions, the lunar and solar equations of motion take the form

$$\ddot{\mathbf{r}} = -n^2 a^3 \frac{\mathbf{r}}{|\mathbf{r}|^3} + n'^2 a'^3 \left[\frac{(\mathbf{r}'-\mathbf{r})}{|\mathbf{r}'-\mathbf{r}|^3} - \frac{\mathbf{r}'}{|\mathbf{r}'|^3}\right], \tag{10.5}$$

$$\ddot{\mathbf{r}}' = -n'^2 a'^3 \frac{\mathbf{r}'}{|\mathbf{r}'|^3}, \tag{10.6}$$

respectively.

Let us set up a conventional Cartesian coordinate system in S which is such that the (apparent) orbit of the Sun about the Earth lies in the x-y plane. This implies that the x-y plane corresponds to the so-called *ecliptic plane*. Accordingly, in S, the Sun appears to orbit the Earth at the mean angular velocity $\boldsymbol{\omega}' = n'\,\mathbf{e}_z$ (assuming that the z-axis points toward the so-called north ecliptic pole), whereas the projection of the Moon onto the ecliptic plane orbits the Earth at the mean angular velocity $\boldsymbol{\omega} = n\,\mathbf{e}_z$.

In the following, for the sake of simplicity, we shall neglect the small eccentricity, $e' = 0.016711$, of the Sun's apparent orbit about the Earth (which is actually the eccentricity of the Earth's orbit about the Sun), and approximate the solar orbit as a *circle*, centered on the Earth. Thus, if x', y', z' are the Cartesian coordinates of the Sun in S then an appropriate solution of the solar equation of motion, (10.6), is

$$x' = a' \cos(n' t), \tag{10.7}$$

$$y' = a' \sin(n' t), \tag{10.8}$$

$$z' = 0. \tag{10.9}$$

10.3 Lunar Equations of Motion

It is convenient to solve the lunar equation of motion, (10.5), in a *geocentric* frame of reference, S_1 (say), which *rotates* with respect to S at the fixed angular velocity $\boldsymbol{\omega}$. Thus, if the lunar orbit were a circle, centered on the Earth, and lying in the ecliptic plane, then the Moon would appear *stationary* in S_1. In fact, the small eccentricity of the lunar orbit, $e = 0.05488$, combined with its slight inclination to the ecliptic plane, $\iota = 5.161^\circ$, causes the Moon to execute a small periodic orbit about the stationary point.

Let x, y, z and x_1, y_1, z_1 be the Cartesian coordinates of the Moon in S and S_1, respectively. It is easily demonstrated that (see Section A.16)

$$x = x_1 \cos(n\,t) - y_1 \sin(n\,t), \tag{10.10}$$

$$y = x_1 \sin(n\,t) + y_1 \cos(n\,t), \tag{10.11}$$

$$z = z_1. \tag{10.12}$$

Moreover, if x_1', y_1', z_1' are the Cartesian components of the Sun in S_1 then (see Section A.5)

$$x_1' = x'\cos(n\,t) + y'\sin(n\,t), \tag{10.13}$$

$$y_1' = -x'\sin(n\,t) + y'\cos(n\,t), \tag{10.14}$$

$$z_1' = z', \tag{10.15}$$

giving

$$x_1' = a'\cos[(n-n')\,t], \tag{10.16}$$

$$y_1' = -a'\sin[(n-n')\,t], \tag{10.17}$$

$$z_1' = 0, \tag{10.18}$$

where use has been made of Equations (10.7)–(10.9).

Now, in the rotating frame S_1, the lunar equation of motion (10.5) transforms to (see Section 6.2)

$$\ddot{\mathbf{r}} + 2\,\boldsymbol{\omega}\times\dot{\mathbf{r}} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) = -n^2\,a^3\,\frac{\mathbf{r}}{|\mathbf{r}|^3} + n'^2\,a'^3\left[\frac{(\mathbf{r}'-\mathbf{r})}{|\mathbf{r}'-\mathbf{r}|^3} - \frac{\mathbf{r}'}{|\mathbf{r}'|^3}\right], \tag{10.19}$$

where $\dot{} \equiv d/dt$. Furthermore, expanding the final term on the right-hand side of (10.19) to lowest order in the small parameter $a/a' = 0.00257$, we obtain

$$\ddot{\mathbf{r}} + 2\,\boldsymbol{\omega}\times\dot{\mathbf{r}} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) \simeq -n^2\,a^3\,\frac{\mathbf{r}}{|\mathbf{r}|^3} + \frac{n'^2\,a'^3}{|\mathbf{r}'|^3}\left[\frac{(3\,\mathbf{r}\cdot\mathbf{r}')\,\mathbf{r}'}{|\mathbf{r}'|^2} - \mathbf{r}\right]. \tag{10.20}$$

When written in terms of Cartesian coordinates, the above equation yields

$$\ddot{x}_1 - 2\,n\,\dot{y}_1 - \left(n^2 + n'^2/2\right)x_1 \simeq -n^2\,a^3\,\frac{x_1}{r^3} + \frac{3}{2}\,n'^2\cos[2\,(n-n')\,t]\,x_1 - \frac{3}{2}\,n'^2\sin[2\,(n-n')\,t]\,y_1, \tag{10.21}$$

$$\ddot{y}_1 + 2\,n\,\dot{x}_1 - \left(n^2 + n'^2/2\right)y_1 \simeq -n^2\,a^3\,\frac{y_1}{r^3} - \frac{3}{2}\,n'^2\sin[2\,(n-n')\,t]\,x_1 - \frac{3}{2}\,n'^2\cos[2\,(n-n')\,t]\,y_1, \tag{10.22}$$

$$\ddot{z}_1 + n'^2\,z_1 \simeq -n^2\,a^3\,\frac{z_1}{r^3}, \tag{10.23}$$

where $r = (x_1^2 + y_1^2 + z_1^2)^{1/2}$, and use has been made of Equations (10.16)–(10.18).

It is convenient, at this stage, to normalize all lengths to a, and all times to n^{-1}. Accordingly, let

$$X = x_1/a, \tag{10.24}$$

$$Y = y_1/a, \tag{10.25}$$

$$Z = z_1/a, \tag{10.26}$$

and $r/a = R = (X^2 + Y^2 + Z^2)^{1/2}$, and $T = n\,t$. In normalized form, Equations (10.21)–(10.23) become

$$\begin{aligned}\ddot{X} - 2\,\dot{Y} - (1 + m^2/2)\,X &\simeq -\frac{X}{R^3} + \frac{3}{2}\,m^2\,\cos[2\,(1-m)\,T]\,X \\ &\quad -\frac{3}{2}\,m^2\,\sin[2\,(1-m)\,T]\,Y, \qquad (10.27)\\ \ddot{Y} + 2\,\dot{X} - (1 + m^2/2)\,Y &\simeq -\frac{Y}{R^3} - \frac{3}{2}\,m^2\,\sin[2\,(1-m)\,T]\,X \\ &\quad -\frac{3}{2}\,m^2\,\cos[2\,(1-m)\,T]\,Y, \qquad (10.28)\\ \ddot{Z} + m^2\,Z &\simeq -\frac{Z}{R^3}, \qquad (10.29)\end{aligned}$$

respectively, where $m = n'/n = 0.07480$ is a measure of the perturbing influence of the Sun on the lunar orbit. Here, $\ddot{} \equiv d^2/dT^2$ and $\dot{} \equiv d/dT$.

Finally, let us write

$$X = X_0 + \delta X, \qquad (10.30)$$

$$Y = \delta Y, \qquad (10.31)$$

$$Z = \delta Z, \qquad (10.32)$$

where $X_0 = (1 + m^2/2)^{-1/3}$, and $|\delta X|$, $|\delta Y|$, $|\delta Z| \ll X_0$. Thus, if the lunar orbit were a circle, centered on the Earth, and lying in the ecliptic plane, then, in the rotating frame S_1, the Moon would appear stationary at the point $(X_0, 0, 0)$. Expanding Equations (10.27)–(10.29) to *second-order* in δX, δY, δZ, and neglecting terms of order m^4 and $m^2\,\delta X^2$, *etc.*, we obtain

$$\begin{aligned}\delta\ddot{X} - 2\,\delta\dot{Y} - 3\,(1 + m^2/2)\,\delta X &\simeq \frac{3}{2}\,m^2\,\cos[2\,(1-m)\,T] + \frac{3}{2}\,m^2\,\cos[2\,(1-m)\,T]\,\delta X \\ &\quad -\frac{3}{2}\,m^2\,\sin[2\,(1-m)\,T]\,\delta Y - 3\,\delta X^2 + \frac{3}{2}\,(\delta Y^2 + \delta Z^2), \qquad (10.33)\\ \delta\ddot{Y} + 2\,\delta\dot{X} &\simeq -\frac{3}{2}\,m^2\,\sin[2\,(1-m)\,T] - \frac{3}{2}\,m^2\,\sin[2\,(1-m)\,T]\,\delta X \\ &\quad -\frac{3}{2}\,m^2\,\cos[2\,(1-m)\,T]\,\delta Y + 3\,\delta X\,\delta Y, \qquad (10.34)\\ \delta\ddot{Z} + (1 + 3\,m^2/2)\,\delta Z &\simeq 3\,\delta X\,\delta Z. \qquad (10.35)\end{aligned}$$

Now, once the above three equations have been solved for δX, δY, and δZ, the Cartesian coordinates, x, y, z, of the Moon in the non-rotating geocentric frame S are obtained from Equations (10.10)–(10.12), (10.24)–(10.26), and (10.30)–(10.32). However, it is more

convenient to write $x = r\cos\theta$, $y = r\sin\theta$, and $z = r\sin\beta$, where r is the *radial distance* between the Earth and Moon, and θ and β are termed the Moon's geocentric (*i.e.*, centered on the Earth) *ecliptic longitude* and *ecliptic latitude*, respectively. Moreover, it is easily seen that, to second-order in δX, δY, δZ, and neglecting terms of order m^4,

$$\frac{r}{a} - 1 + \frac{m^2}{6} \simeq \delta X + \frac{1}{2}\,\delta Y^2 + \frac{1}{2}\,\delta Z^2, \tag{10.36}$$

$$\theta - n\,t \simeq \delta Y - \delta X\,\delta Y, \tag{10.37}$$

$$\beta \simeq \delta Z - \delta X\,\delta Z. \tag{10.38}$$

10.4 Unperturbed Lunar Motion

Let us, first of all, neglect the perturbing influence of the Sun on the Moon's orbit by setting $m = 0$ in the lunar equations of motion (10.33)–(10.35). For the sake of simplicity, let us also neglect nonlinear effects in these equations by setting $\delta X^2 = \delta Y^2 = \delta Z^2 = \delta X\,\delta Y = \delta X\,\delta Z = 0$. In this case, the equations reduce to

$$\delta\ddot{X} - 2\,\delta\dot{Y} - 3\,\delta X \simeq 0, \tag{10.39}$$

$$\delta\ddot{Y} + 2\,\delta\dot{X} \simeq 0, \tag{10.40}$$

$$\delta\ddot{Z} + \delta Z \simeq 0. \tag{10.41}$$

By inspection, appropriate solutions are

$$\delta X \simeq -e\,\cos(T - \alpha_0), \tag{10.42}$$

$$\delta Y \simeq 2\,e\,\sin(T - \alpha_0), \tag{10.43}$$

$$\delta Z \simeq i\,\sin(T - \gamma_0), \tag{10.44}$$

where e, α_0, i, and γ_0 are arbitrary constants. Recalling that $T = n\,t$, it follows from (10.36)–(10.38) that

$$r \simeq a\,[1 - e\,\cos(n\,t - \alpha_0)]\,, \tag{10.45}$$

$$\theta \simeq n\,t + 2\,e\,\sin(n\,t - \alpha_0), \tag{10.46}$$

$$\beta \simeq \iota\,\sin(n\,t - \gamma_0). \tag{10.47}$$

However, Equations (10.45) and (10.46) are simply first-order (in e) approximations to the familiar Keplerian laws (see Chapter 4)

$$r = \frac{a\,(1 - e^2)}{1 + e\,\cos(\theta - \alpha_0)}, \tag{10.48}$$

$$r^2\,\dot{\theta} = (1 - e^2)^{1/2}\,n\,a^2, \tag{10.49}$$

where $\dot{} \equiv d/dt$. Of course, the above two laws describe a body which executes an *elliptical* orbit, confocal with the Earth, of major radius a, mean angular velocity n, and eccentricity e, such that the radius vector connecting the body to the Earth sweeps out equal areas in equal time intervals. We conclude, unsurprisingly, that the unperturbed lunar orbit is a Keplerian ellipse. Note that the lunar *perigee* lies at the fixed ecliptic longitude $\theta = \alpha_0$. Equation (10.47) is the first-order approximation to

$$\beta = \iota \sin(\theta - \gamma_0). \tag{10.50}$$

This expression implies that the unperturbed lunar orbit is *co-planar*, but is *inclined* at an angle ι to the ecliptic plane. Moreover, the *ascending node* lies at the fixed ecliptic longitude $\theta = \gamma_0$. Incidentally, the neglect of nonlinear terms in Equations (10.39)–(10.41) is only valid as long as $e, \iota \ll 1$: *i.e.*, provided that the unperturbed lunar orbit is only slightly elliptical, and slightly inclined to the ecliptic plane. In fact, the observed values of e and ι are 0.05488 and 0.09008 radians, respectively, so this is indeed the case.

10.5 Perturbed Lunar Motion

The perturbed nonlinear lunar equations of motion, (10.33)–(10.35), take the general form

$$\delta\ddot{X} - 2\,\delta\dot{Y} - 3\,(1 + m^2/2)\,\delta X \simeq R_X, \tag{10.51}$$

$$\delta\ddot{Y} + 2\,\delta\dot{X} \simeq R_Y, \tag{10.52}$$

$$\delta\ddot{Z} + (1 + 3\,m^2/2)\,\delta Z \simeq R_Z, \tag{10.53}$$

where

$$R_X = a_0 + \sum_{j>0} a_j \cos(\omega_j\,T - \alpha_j), \tag{10.54}$$

$$R_Y = \sum_{j>0} b_j \sin(\omega_j\,T - \alpha_j), \tag{10.55}$$

$$R_Z = \sum_{j>0} c_j \sin(\Omega_j\,T - \gamma_j). \tag{10.56}$$

Let us search for solutions of the general form

$$\delta X = x_0 + \sum_{j>0} x_j \cos(\omega_j\,T - \alpha_j), \tag{10.57}$$

$$\delta Y = \sum_{j>0} y_j \sin(\omega_j\,T - \alpha_j), \tag{10.58}$$

$$\delta Z = \sum_{j>0} z_j \sin(\Omega_j\,T - \gamma_j). \tag{10.59}$$

j	ω_j	α_j	Ω_j	γ_j
1	$1 + c\,m^2$	α_0	$1 + g\,m^2$	γ_0
2	$2\,(1 + c\,m^2)$	$2\,\alpha_0$	$(c - g)\,m^2$	$\alpha_0 - \gamma_0$
3	$2\,(1 + g\,m^2)$	$2\,\gamma_0$	$2 + (c + g)\,m^2$	$\alpha_0 + \gamma_0$
4	$2 - 2\,m$	0		
5	$1 - 2\,m - c\,m^2$	$-\alpha_0$	$1 - 2\,m - g\,m^2$	$-\gamma_0$

Table 10.1: *Angular frequencies and phase-shifts associated with the principal periodic driving terms appearing in the perturbed nonlinear lunar equations of motion.*

Substituting expressions (10.54)–(10.59) into Equations (10.51)–(10.53), it is easily demonstrated that

$$x_0 = -\frac{a_0}{3\,(1 + m^2/2)}, \tag{10.60}$$

$$x_j = \frac{\omega_j\,a_j - 2\,b_j}{\omega_j\,(1 - 3\,m^2/2 - \omega_j^2)}, \tag{10.61}$$

$$y_j = \frac{(\omega_j^2 + 3 + 3\,m^2/2)\,b_j - 2\,\omega_j\,a_j}{\omega_j^2\,(1 - 3\,m^2/2 - \omega_j^2)}, \tag{10.62}$$

$$z_j = \frac{c_j}{1 + 3\,m^2/2 - \Omega_j^2}, \tag{10.63}$$

where $j > 0$.

The angular frequencies, ω_j, Ω_j, and phase shifts, α_j, γ_j, of the principal periodic driving terms that appear on the right-hand sides of the perturbed nonlinear lunar equations of motion, (10.51)–(10.53), are specified in Table 10.1. Here, c and g are, as yet, unspecified $\mathcal{O}(1)$ constants associated with the precession of the lunar perigee, and the regression of the ascending node, respectively. Note that ω_1 and Ω_1 are the frequencies of the Moon's unforced motion in ecliptic longitude and latitude, respectively. Moreover, ω_4 is the forcing frequency associated with the perturbing influence of the Sun. All other frequencies appearing in Table 10.1 are combinations of these three fundamental frequencies. In fact, $\omega_2 = 2\,\omega_1$, $\omega_3 = 2\,\Omega_1$, $\omega_5 = \omega_4 - \omega_1$, $\Omega_2 = \omega_1 - \Omega_1$, $\Omega_3 = \omega_1 + \Omega_1$, and $\Omega_5 = \omega_4 - \Omega_1$. Note that there is no Ω_4.

Now, a comparison of Equations (10.33)–(10.35), (10.51)–(10.53), and Table 10.1 reveals that

$$\begin{aligned} R_X &= \frac{3}{2}\,m^2\,\cos(\omega_4\,T - \alpha_4) + \frac{3}{2}\,m^2\,\cos(\omega_4\,T - \alpha_4)\,\delta X \\ &\quad -\frac{3}{2}\,m^2\,\sin(\omega_4\,T - \alpha_4)\,\delta Y - 3\,\delta X^2 + \frac{3}{2}\,(\delta Y^2 + \delta Z^2), \end{aligned} \tag{10.64}$$

$$R_Y = -\frac{3}{2}\,m^2\,\sin(\omega_4\,T - \alpha_4) - \frac{3}{2}\,m^2\,\sin(\omega_4\,T - \alpha_4)\,\delta X$$

j	a_j	b_j	c_j
0	$\frac{3}{2}\,e^2 + \frac{3}{4}\,\iota^2$		
1	$\frac{3}{4}\,m^2\,x_5 - \frac{3}{4}\,m^2\,y_5 - 3\,x_4\,x_5 + \frac{3}{2}\,y_4\,y_5$	$-\frac{3}{4}\,m^2\,x_5 + \frac{3}{4}\,m^2\,y_5 + \frac{3}{2}\,y_4\,x_5 - \frac{3}{2}\,y_5\,x_4$	$-\frac{3}{2}\,x_4\,z_5$
2	$-\frac{9}{2}\,e^2$	$-3\,e^2$	$\frac{3}{2}\,e\,\iota$
3	$-\frac{3}{4}\,\iota^2$	0	$-\frac{3}{2}\,e\,\iota$
4	$\frac{3}{2}\,m^2$	$-\frac{3}{2}\,m^2$	0
5	$-\frac{9}{4}\,m^2\,e + 3\,e\,x_4 + 3\,e\,y_4$	$\frac{9}{4}\,m^2\,e - 3\,e\,x_4 - \frac{3}{2}\,e\,y_4$	$-\frac{3}{2}\,\iota\,x_4$

Table 10.2: *Amplitudes of the periodic driving terms appearing in the perturbed nonlinear lunar equations of motion.*

$$-\frac{3}{2}\,m^2\,\cos(\omega_4\,T - \alpha_4)\,\delta Y + 3\,\delta X\,\delta Y, \tag{10.65}$$

$$R_Z = 3\,\delta X\,\delta Z. \tag{10.66}$$

Substitution of the solutions (10.57)–(10.59) into the above equations, followed by a comparison with expressions (10.54)–(10.56), yields the amplitudes a_j, b_j, and c_j specified in Table 10.2. Note that, in calculating these amplitudes, we have neglected all contributions to the periodic driving terms, appearing in Equations (10.51)–(10.53) which involve cubic, or higher order, combinations of e, ι, m^2, x_j, y_j, and z_j, since we only expanded Equations (10.33)–(10.35) to *second-order* in δX, δY, and δZ.

For $j = 0$, it follows from Equation (10.60) and Table 10.2 that

$$x_0 = -\frac{1}{2}\,e^2 - \frac{1}{4}\,\iota^2. \tag{10.67}$$

For $j = 2$, making the approximation $\omega_2 \simeq 2$ (see Table 10.1), it follows from Equations (10.61), (10.62) and Table 10.2 that

$$x_2 \simeq \frac{1}{2}\,e^2, \tag{10.68}$$

$$y_2 \simeq \frac{1}{4}\,e^2. \tag{10.69}$$

Likewise, making the approximation $\Omega_2 \simeq 0$ (see Table 10.1), it follows from Equation (10.63) and Table 10.2 that

$$z_2 \simeq \frac{3}{2}\,e\,\iota. \tag{10.70}$$

For $j = 3$, making the approximation $\omega_3 \simeq 2$ (see Table 10.1), it follows from Equations (10.61), (10.62) and Table 10.2 that

$$x_3 \simeq \frac{1}{4}\,\iota^2, \tag{10.71}$$

$$y_3 \simeq -\frac{1}{4}\,\iota^2. \tag{10.72}$$

Likewise, making the approximation $\Omega_3 \simeq 2$ (see Table 10.1), it follows from Equation (10.63) and Table 10.2 that

$$z_3 \simeq \frac{1}{2}\, e\, \iota. \tag{10.73}$$

For $j = 4$, making the approximation $\omega_4 \simeq 2$ (see Table 10.1), it follows from Equations (10.61), (10.62) and Table 10.2 that

$$x_4 \simeq -m^2, \tag{10.74}$$

$$y_4 \simeq \frac{11}{8}\, m^2. \tag{10.75}$$

Thus, according to Table 10.2,

$$a_5 \simeq -\frac{9}{8}\, m^2\, e, \tag{10.76}$$

$$b_5 \simeq \frac{51}{16}\, m^2\, e, \tag{10.77}$$

$$c_5 \simeq \frac{3}{2}\, m^2\, \iota. \tag{10.78}$$

For $j = 5$, making the approximation $\omega_5 \simeq 1 - 2\, m$ (see Table 10.1), it follows from Equations (10.61), (10.62), (10.76), and (10.77) that

$$x_5 \simeq -\frac{15}{8}\, m\, e, \tag{10.79}$$

$$y_5 \simeq \frac{15}{4}\, m\, e. \tag{10.80}$$

Likewise, making the approximation $\Omega_5 \simeq 1 - 2\, m$ (see Table 10.1), it follows from Equations (10.63) and (10.78) that

$$z_5 \simeq \frac{3}{8}\, m\, \iota. \tag{10.81}$$

Thus, according to Table 10.2,

$$a_1 = -\frac{135}{64}\, m^3\, e, \tag{10.82}$$

$$b_1 = \frac{765}{128}\, m^3\, e, \tag{10.83}$$

$$c_1 = \frac{9}{16}\, m^3\, \iota. \tag{10.84}$$

Finally, for $j = 1$, by analogy with Equations (10.42)–(10.44), we expect

$$a_1 = -e, \tag{10.85}$$

$$b_1 = 2\, e, \tag{10.86}$$

$$c_1 = \iota. \tag{10.87}$$

Thus, since $\omega_1 = 1 + c\,m^2$ (see Table 10.1), it follows from Equations (10.61), (10.82), (10.83), and (10.85) that

$$-e \simeq \frac{-(225/16)\,m^3\,e}{-(3/2)\,m^2 - 2\,c\,m^2}, \tag{10.88}$$

which yields

$$c \simeq -\frac{3}{4} - \frac{225}{32}\,m + \mathcal{O}(m^2). \tag{10.89}$$

Likewise, since $\Omega_1 = 1 + g\,m^2$ (see Table 10.1), it follows from Equations (10.63), (10.84), and (10.87) that

$$\iota \simeq \frac{(9/16)\,m^3\,\iota}{(3/2)\,m^2 - 2\,g\,m^2}, \tag{10.90}$$

which yields

$$g \simeq \frac{3}{4} + \frac{9}{32}\,m + \mathcal{O}(m^2). \tag{10.91}$$

According to the above analysis, our final expressions for δX, δY, and δZ are

$$\begin{aligned} \delta X &= -\frac{1}{2}\,e^2 - \frac{1}{4}\,\iota^2 - e\,\cos[(1 + c\,m^2)\,T - \alpha_0] + \frac{1}{2}\,e^2\,\cos[2\,(1 + c\,m^2)\,T - 2\,\alpha_0] \\ &\quad + \frac{1}{4}\,\iota^2\,\cos[2\,(1 + g\,m^2)\,T - 2\,\gamma_0] - m^2\,\cos[2\,(1 - m)\,T] \\ &\quad - \frac{15}{8}\,m\,e\,\cos[(1 - 2\,m - c\,m^2)\,T + \alpha_0], \end{aligned} \tag{10.92}$$

$$\begin{aligned} \delta Y &= 2\,e\,\sin[(1 + c\,m^2)\,T - \alpha_0] + \frac{1}{4}\,e^2\,\sin[2\,(1 + c\,m^2)\,T - 2\,\alpha_0] \\ &\quad - \frac{1}{4}\,\iota^2\,\sin[2\,(1 + g\,m^2)\,T - 2\,\gamma_0] + \frac{11}{8}\,m^2\,\cos[2\,(1 - m)\,T] \\ &\quad + \frac{15}{4}\,m\,e\,\sin[(1 - 2\,m - c\,m^2)\,T + \alpha_0], \end{aligned} \tag{10.93}$$

$$\begin{aligned} \delta Z &= \iota\,\sin[(1 + g\,m^2)\,T - \gamma_0] + \frac{3}{2}\,e\,\iota\,\sin[(c - g)\,m^2\,T - \alpha_0 + \gamma_0] \\ &\quad + \frac{1}{2}\,e\,\iota\,\sin[(2 + c\,m^2 + g\,m^2)\,T - \alpha_0 - \gamma_0] \\ &\quad + \frac{3}{8}\,m\,\iota\,\sin[(1 - 2\,m - g\,m^2)\,T + \gamma_0]. \end{aligned} \tag{10.94}$$

Thus, making use of Equations (10.36)–(10.38), we find that

$$\begin{aligned} \frac{r}{a} &= 1 - e\,\cos[(1 + c\,m^2)\,T - \alpha_0] + \frac{1}{2}\,e^2 - \frac{1}{6}\,m^2 - \frac{1}{2}\,e^2\,\cos[2\,(1 + c\,m^2)\,T - 2\,\alpha_0] \\ &\quad - m^2\,\cos[2\,(1 - m)\,T] - \frac{15}{8}\,m\,e\,\cos[(1 - 2\,m - c\,m^2)\,T + \alpha_0], \end{aligned} \tag{10.95}$$

$$\begin{aligned}\theta &= T + 2\,e\,\sin[(1+c\,m^2)\,T - \alpha_0] + \frac{5}{4}\,e^2\,\sin[2\,(1+c\,m^2)\,T - 2\,\alpha_0]\\ &\quad -\frac{1}{4}\,\iota^2\,\sin[2\,(1+g\,m^2)\,T - 2\,\gamma_0] + \frac{11}{8}\,m^2\,\sin[2\,(1-m)\,T]\\ &\quad +\frac{15}{4}\,m\,e\,\sin[(1-2\,m-c\,m^2)\,T + \alpha_0], \end{aligned} \tag{10.96}$$

$$\begin{aligned}\beta &= \iota\,\sin[(1+g\,m^2)\,T - \gamma_0] + e\,\iota\,\sin[(c-g)\,m^2\,T - \alpha_0 + \gamma_0]\\ &\quad + e\,\iota\,\sin[(2+c\,m^2+g\,m^2)\,T - \alpha_0 - \gamma_0]\\ &\quad +\frac{3}{8}\,m\,\iota\,\sin[(1-2\,m-g\,m^2)\,T + \gamma_0]. \end{aligned} \tag{10.97}$$

The above expressions are accurate up to *second-order* in the small parameters e, ι, and m.

10.6 Description of Lunar Motion

In order to better understand the perturbed lunar motion derived in the previous section, it is helpful to introduce the concept of the *mean moon*. This is an imaginary body which orbits the Earth, in the ecliptic plane, at a *steady* angular velocity that is equal to the Moon's mean orbital angular velocity, n. Likewise, the *mean sun* is a second imaginary body which orbits the Earth, in the ecliptic plane, at a *steady* angular velocity that is equal to the Sun's mean orbital angular velocity, n'. Thus, the ecliptic longitudes of the mean moon and the mean sun are

$$\bar{\theta} = n\,t, \tag{10.98}$$

$$\bar{\theta}' = n'\,t, \tag{10.99}$$

respectively. Here, for the sake of simplicity, and also for the sake of consistency with our previous analysis, we have assumed that both objects are located at ecliptic longitude 0° at time $t = 0$.

Now, from Equation (10.95), to first-order in small parameters, the lunar perigee corresponds to $(1 + c\,m^2)\,n\,t - \alpha_0 = j\,2\pi$, where j is an integer. However, this condition can also be written $\bar{\theta} = \alpha$, where

$$\alpha = \alpha_0 + \alpha_1\,n'\,t, \tag{10.100}$$

and, making use of Equation (10.89), together with the definition $m = n'/n$,

$$\alpha_1 = \frac{3}{4}\,m + \frac{255}{32}\,m^2 + \mathcal{O}(m^3). \tag{10.101}$$

Thus, we can identify α as the mean ecliptic longitude of the perigee. Moreover, according to Equation (10.100), the perigee *precesses* (*i.e.*, its longitude increases in time) at the mean rate of $360\,\alpha_1$ degrees per year. (Of course, a year corresponds to $\Delta t = 2\pi/n'$.) Furthermore, it is clear that this precession is entirely due to the perturbing influence of the

Sun, since it only depends on the parameter m, which is a measure of this influence. Given that $m = 0.07480$, we find that the perigee advances by 34.36° degrees per year. Hence, we predict that the perigee completes a full circuit about the Earth every $1/\alpha_1 = 10.5$ years. In fact, the lunar perigee completes a full circuit every 8.85 years. Our prediction is somewhat inaccurate because our previous analysis neglected $\mathcal{O}(m^2)$, and smaller, contributions to the parameter c [see Equation (10.89)], and these turn out to be significant.

From Equation (10.97), to first-order in small parameters, the Moon passes through its ascending node when $(1 + g\,m^2)\,n\,t - \gamma_0 = \gamma_0 + j\,2\pi$, where j is an integer. However, this condition can also be written $\bar{\theta} = \gamma$, where

$$\gamma = \gamma_0 - \gamma_1\,n'\,t, \tag{10.102}$$

and, making use of Equation (10.91),

$$\gamma_1 = \frac{3}{4}\,m - \frac{9}{32}\,m^2 + \mathcal{O}(m^3). \tag{10.103}$$

Thus, we can identify γ as the mean ecliptic longitude of the ascending node. Moreover, according to Equation (10.102), the ascending node *regresses* (*i.e.*, its longitude decreases in time) at the mean rate of $360\,\gamma_1$ degrees per year. As before, it is clear that this regression is entirely due to the perturbing influence of the Sun. Moreover, we find that the ascending node retreats by 19.63° degrees per year. Hence, we predict that the ascending node completes a full circuit about the Earth every $1/\gamma_1 = 18.3$ years. In fact, the lunar ascending node completes a full circuit every 18.6 years, so our prediction is fairly accurate.

It is helpful to introduce the lunar *mean anomaly*,

$$M = \bar{\theta} - \alpha, \tag{10.104}$$

which is defined as the angular distance (in longitude) between the mean Moon and the perigee. It is also helpful to introduce the lunar *mean argument of latitude*,

$$F = \bar{\theta} - \gamma, \tag{10.105}$$

which is defined as the angular distance (in longitude) between the mean Moon and the ascending node. Finally, it is helpful to introduce the *mean elongation* of the Moon,

$$D = \bar{\theta} - \bar{\theta}', \tag{10.106}$$

which is defined as the difference between the longitudes of the mean Moon and the mean Sun.

When expressed in terms of M, F, and D, our previous expression (10.96) for the true ecliptic longitude of the Moon becomes

$$\theta = \bar{\theta} + \lambda, \tag{10.107}$$

where

$$\lambda = 2\,e\,\sin M + \frac{5}{4}\,e^2\,\sin 2\,M - \frac{1}{4}\,\iota^2\,\sin 2\,F + \frac{11}{8}\,m^2\,\sin 2\,D + \frac{15}{4}\,m\,e\,\sin(2\,D - M) \tag{10.108}$$

is the angular distance (in longitude) between the Moon and the mean Moon. The first three terms on the right-hand side of the above expression are Keplerian (*i.e.*, they are independent of the perturbing action of the Sun). In fact, the first is due to the eccentricity of the lunar orbit (*i.e.*, the fact that the geometric center of the orbit is slightly shifted from the Earth), the second is due to the ellipticity of the orbit (*i.e.*, the fact that the orbit is slightly non-circular), and the third is due to the slight inclination of the orbit to the ecliptic plane. However, the final two terms are caused by the perturbing action of the Sun. In fact, the fourth term corresponds to *variation*, whilst the fifth corresponds to *evection*. Note that variation attains its maximal amplitude around the so-called *octant points*, at which the Moon's disk is either one-quarter or three-quarters illuminated (*i.e.*, when $D = 45^\circ$, 135°, 225°, or 315°). Conversely, the amplitude of variation is zero around the so-called *quadrant points*, at which the Moon's disk is either fully illuminated, half illuminated, or not illuminated at all (*i.e.*, when $D = 0^\circ$, 90°, 180°, or 270°). Evection can be thought of as causing a slight reduction in the eccentricity of the lunar orbit around the times of the new moon and the full moon (*i.e.*, $D = 0^\circ$ and $D = 180^\circ$), and causing a corresponding slight increase in the eccentricity around the times of the first and last quarter moons (*i.e.*, $D = 90^\circ$ and $D = 270^\circ$). This follows because the evection term in Equation (10.108) augments the eccentricity term, $2\,e\,\sin M$, when $\cos 2D = -1$, and reduces the term when $\cos 2D = +1$. The variation and evection terms appearing in expression (10.108) oscillate sinusoidally with periods of half a synodic month,[2] or 14.8 days, and 31.8 days, respectively. These periods are in good agreement with observations. Finally, the amplitudes of the variation and evection terms (calculated using $m = 0.07480$ and $e = 0.05488$) are 1630 and 3218 arc seconds, respectively. However, the observed amplitudes are 2370 and 4586 arc seconds, respectively. It turns out that our expressions for these amplitudes are somewhat inaccurate because, for the sake of simplicity, we have only calculated the lowest order (in m) contributions to them. Recall that we also neglected the slight eccentricity, $e' = 0.016711$, of the Sun's apparent orbit about the Earth in our calculation. In fact, the eccentricity of the solar orbit gives rise to a small addition term $-3\,m\,e'\,\sin M'$ on the right-hand side of (10.108), where M' is the Sun's mean anomaly. This term, which is known as the *annual equation*, oscillates with a period of a solar year, and has an amplitude of 772 arc seconds.

When expressed in terms of D and F our previous expression (10.97) for the ecliptic latitude of the Moon becomes

$$\beta = \iota\,\sin(F + \lambda) + \frac{3}{8}\,m\,\iota\,\sin(2\,D - F). \tag{10.109}$$

[2]A synodic month, which is 29.53 days, is the mean period between successive new moons.

The first term on the right-hand side of this expression is Keplerian (*i.e.*, it is independent of the perturbing influence of the Sun). However, the final term, which is known as *evection in latitude*, is due to the Sun's action. Evection in latitude can be thought of as causing a slight increase in the inclination of the lunar orbit to the ecliptic at the times of the first and last quarter moons, and a slight decrease at the times of the new moon and the full moon. The evection in latitude term oscillates sinusoidally with a period of 32.3 days, and has an amplitude of 521 arc seconds. This period is in good agreement with observations. However, the observed amplitude of the evection in latitude term is 624 arc seconds. As before, our expression for the amplitude is somewhat inaccurate because we have only calculated the lowest order (in m) contribution.

11 Orbital Perturbation Theory

11.1 Introduction

The two-body orbit theory described in Chapter 4 neglects the gravitational interactions between the various different planets in the Solar System, whilst retaining those between each individual planet and the Sun. This is an excellent first approximation, since the former interactions are much weaker than the latter, as a consequence of the small masses of the Planets relative to the Sun (see Table 4.1). Nevertheless, interplanetary gravitational interactions do have a profound influence on planetary orbits when integrated over long periods of time. The aim of this chapter is to examine this influence in a systematic fashion using a branch of celestial mechanics known as *perturbation theory*.

11.2 Disturbing Function

For the moment, let us consider a simplified Solar System that consists of the Sun and *two* planets—see Figure 11.1. Let the Sun be of mass M, and position vector $\mathbf{R}_s$. Likewise, let the two planets have masses m and m', and position vectors $\mathbf{R}$ and $\mathbf{R}'$, respectively. Here, we are assuming that $m,\ m' \ll M$. Let $\mathbf{r} = \mathbf{R} - \mathbf{R}_s$ and $\mathbf{r}' = \mathbf{R}' - \mathbf{R}_s$, be the position vector of each planet relative to the Sun. Without loss of generality, we can assume that $r' > r$.

Now, the equations of motion of the various elements of our simplified Solar System are

$$M\,\ddot{\mathbf{R}}_s = G\,M\,m\,\frac{\mathbf{r}}{r^3} + G\,M\,m'\,\frac{\mathbf{r}'}{r'^3}, \tag{11.1}$$

$$m\,\ddot{\mathbf{R}} = G\,m\,m'\,\frac{\mathbf{r}'-\mathbf{r}}{|\mathbf{r}'-\mathbf{r}|^3} - G\,m\,M\,\frac{\mathbf{r}}{r^3}, \tag{11.2}$$

$$m'\,\ddot{\mathbf{R}}' = G\,m'\,m\,\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} - G\,m'\,M\,\frac{\mathbf{r}'}{r'^3}, \tag{11.3}$$

It thus follows that

$$\ddot{\mathbf{r}} + \mu\,\frac{\mathbf{r}}{r^3} = G\,m'\left(\frac{\mathbf{r}'-\mathbf{r}}{|\mathbf{r}'-\mathbf{r}|^3} - \frac{\mathbf{r}'}{r'^3}\right), \tag{11.4}$$

$$\ddot{\mathbf{r}}' + \mu'\,\frac{\mathbf{r}'}{r'^3} = G\,m\left(\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} - \frac{\mathbf{r}}{r^3}\right), \tag{11.5}$$

where $\mu = G\,(M+m)$, and $\mu' = G\,(M+m')$. The right-hand sides of the above equations specify the interplanetary interaction forces that were neglected in our previous analysis.

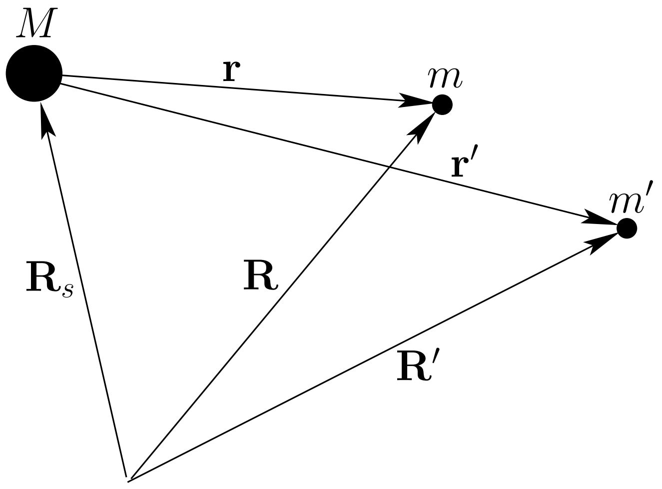


Figure 11.1: *A simplified model of the Solar System.*

These right-hand sides can be conveniently expressed as the gradients of potentials: *i.e.*,

$$\ddot{\mathbf{r}} + \mu\,\frac{\mathbf{r}}{r^3} = \nabla\mathcal{R}, \tag{11.6}$$

$$\ddot{\mathbf{r}}' + \mu'\,\frac{\mathbf{r}'}{r'^3} = \nabla'\mathcal{R}', \tag{11.7}$$

where

$$\mathcal{R}(\mathbf{r},\mathbf{r}') = \tilde{\mu}'\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{\mathbf{r}\cdot\mathbf{r}'}{r'^3}\right), \tag{11.8}$$

$$\mathcal{R}'(\mathbf{r},\mathbf{r}') = \tilde{\mu}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{\mathbf{r}\cdot\mathbf{r}'}{r^3}\right), \tag{11.9}$$

with $\tilde{\mu} = G\,m$, and $\tilde{\mu}' = G\,m'$. Here, $\mathcal{R}(\mathbf{r},\mathbf{r}')$ and $\mathcal{R}'(\mathbf{r},\mathbf{r}')$ are termed *disturbing functions*. Moreover, ∇ and ∇' are the gradient operators involving the unprimed and primed coordinates, respectively.

11.3 Osculating Orbital Elements

Consider Equation (11.6), which can be written

$$\ddot{x} + \mu\,\frac{x}{r^3} = \frac{\partial\mathcal{R}}{\partial x}, \tag{11.10}$$

$$\ddot{y} + \mu\,\frac{y}{r^3} = \frac{\partial\mathcal{R}}{\partial y}, \tag{11.11}$$

$$\ddot{z} + \mu\,\frac{z}{r^3} = \frac{\partial\mathcal{R}}{\partial z}. \tag{11.12}$$

If the right-hand sides of the above equations are set to zero then the problem reduces to that considered in Chapter 4, and we thus have solutions which can be written in the form

$$
\begin{aligned}
x &= f_1(c_1, c_2, c_3, c_4, c_5, c_6, t), && (11.13)\\
y &= f_2(c_1, c_2, c_3, c_4, c_5, c_6, t), && (11.14)\\
z &= f_3(c_1, c_2, c_3, c_4, c_5, c_6, t), && (11.15)\\
\dot{x} &= g_1(c_1, c_2, c_3, c_4, c_5, c_6, t), && (11.16)\\
\dot{y} &= g_2(c_1, c_2, c_3, c_4, c_5, c_6, t), && (11.17)\\
\dot{z} &= g_3(c_1, c_2, c_3, c_4, c_5, c_6, t). && (11.18)
\end{aligned}
$$

Here, $c_1, \cdots, c_6$ are the six orbital elements of the first planet (see Section 4.10), which can be thought of as constants of its motion. It follows that

$$g_k = \frac{\partial f_k}{\partial t}, \tag{11.19}$$

for $k = 1, 2, 3$.

Let us now take the right-hand sides of Equations (11.10)–(11.12) into account. The orbital elements, $c_1, \ldots, c_6$, are no longer constants of the motion. However, we would expect them to be relatively *slowly varying* functions of time, given the relative smallness of the right-hand sides. Let us derive evolution equations for these so-called *osculating orbital elements*.

According to Equation (11.13), we have

$$\frac{dx}{dt} = \frac{\partial f_1}{\partial t} + \sum_{k=1,6} \frac{\partial f_1}{\partial c_k}\frac{dc_k}{dt}. \tag{11.20}$$

If this expression, and the analogous expressions for dy/dt and dz/dt, were differentiated in time, and the results substituted into Equations (11.10)–(11.12), then we would obtain three time evolution equations for the six variables $c_1, \cdots, c_6$. In order to make the problem definite, three additional conditions must be introduced into the problem. It is convenient to choose

$$\sum_{k=1,6} \frac{\partial f_l}{\partial c_k}\frac{dc_k}{dt} = 0, \tag{11.21}$$

for $l = 1, 2, 3$. Hence, it follows that

$$
\begin{aligned}
\frac{dx}{dt} &= \frac{\partial f_1}{\partial t} = g_1, && (11.22)\\
\frac{dy}{dt} &= \frac{\partial f_2}{\partial t} = g_2, && (11.23)\\
\frac{dz}{dt} &= \frac{\partial f_3}{\partial t} = g_3. && (11.24)
\end{aligned}
$$

Differentiation of the above equations with respect to time yields

$$\frac{d^2x}{dt^2} = \frac{\partial^2 f_1}{\partial t^2} + \sum_{k=1,6} \frac{\partial g_1}{\partial c_k}\frac{dc_k}{dt}, \tag{11.25}$$

$$\frac{d^2y}{dt^2} = \frac{\partial^2 f_2}{\partial t^2} + \sum_{k=1,6} \frac{\partial g_2}{\partial c_k}\frac{dc_k}{dt}, \tag{11.26}$$

$$\frac{d^2z}{dt^2} = \frac{\partial^2 f_3}{\partial t^2} + \sum_{k=1,6} \frac{\partial g_3}{\partial c_k}\frac{dc_k}{dt}. \tag{11.27}$$

Substitution into Equations (11.10)–(11.12) gives

$$\frac{\partial^2 f_1}{\partial t^2} + \mu\frac{f_1}{r^3} + \sum_{k=1,6} \frac{\partial g_1}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial x}, \tag{11.28}$$

$$\frac{\partial^2 f_2}{\partial t^2} + \mu\frac{f_2}{r^3} + \sum_{k=1,6} \frac{\partial g_2}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial y}, \tag{11.29}$$

$$\frac{\partial^2 f_3}{\partial t^2} + \mu\frac{f_3}{r^3} + \sum_{k=1,6} \frac{\partial g_3}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial z}, \tag{11.30}$$

where $r = (f_1^2 + f_2^2 + f_3^2)^{1/2}$. Since f_1, f_2, and f_3 are the respective solutions to the above equations when the right-hand sides are zero, and the orbital elements are thus constants, it follows that the first two terms in each equation cancel one another. Hence, writing f_1 as x, and g_1 as $\dot{x}$, *etc.*, Equations (11.21) and (11.28)–(11.30) yield

$$\sum_{k=1,6} \frac{\partial x}{\partial c_k}\frac{dc_k}{dt} = 0, \tag{11.31}$$

$$\sum_{k=1,6} \frac{\partial y}{\partial c_k}\frac{dc_k}{dt} = 0, \tag{11.32}$$

$$\sum_{k=1,6} \frac{\partial z}{\partial c_k}\frac{dc_k}{dt} = 0, \tag{11.33}$$

$$\sum_{k=1,6} \frac{\partial \dot{x}}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial x}, \tag{11.34}$$

$$\sum_{k=1,6} \frac{\partial \dot{y}}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial y}, \tag{11.35}$$

$$\sum_{k=1,6} \frac{\partial \dot{z}}{\partial c_k}\frac{dc_k}{dt} = \frac{\partial \mathcal{R}}{\partial z}. \tag{11.36}$$

These six equations are equivalent to the three original equations of motion, (11.10)–(11.12).

11.4 Lagrange Brackets

Six new equations can be derived from Equations (11.31)–(11.36) by multiplying them successively by $-\partial\dot{x}/\partial c_j$, $-\partial\dot{y}/\partial c_j$, $-\partial\dot{z}/\partial c_j$, $\partial x/\partial c_j$, $\partial y/\partial c_j$, $\partial z/\partial c_j$, and summing the resulting equations. The right-hand sides of the new equations are

$$\frac{\partial\mathcal{R}}{\partial x}\frac{\partial x}{\partial c_j}+\frac{\partial\mathcal{R}}{\partial y}\frac{\partial y}{\partial c_j}+\frac{\partial\mathcal{R}}{\partial z}\frac{\partial z}{\partial c_j}\equiv\frac{\partial\mathcal{R}}{\partial c_j}. \tag{11.37}$$

The new equations can be written in a more compact form via the introduction of *Lagrange brackets,* which are defined

$$[c_j,c_k]=\sum_{l=1,3}\left(\frac{\partial x_l}{\partial c_j}\frac{\partial\dot{x}_l}{\partial c_k}-\frac{\partial x_l}{\partial c_k}\frac{\partial\dot{x}_l}{\partial c_j}\right), \tag{11.38}$$

where $x_i\equiv x$, $x_2\equiv y$, and $x_3\equiv z$. Thus, the new equations become

$$\sum_{k=1,6}[c_j,c_k]\frac{dc_k}{dt}=\frac{\partial\mathcal{R}}{\partial c_j}, \tag{11.39}$$

for $j=1,6$. Note, incidentally, that

$$[c_j,c_j] = 0, \tag{11.40}$$

$$[c_j,c_k] = -[c_k,c_j]. \tag{11.41}$$

Let

$$[p,q]=\sum_{l=1,3}\left(\frac{\partial x_l}{\partial p}\frac{\partial\dot{x}_l}{\partial q}-\frac{\partial x_l}{\partial q}\frac{\partial\dot{x}_l}{\partial p}\right), \tag{11.42}$$

where p and q are two orbital elements. It follows that

$$\frac{\partial}{\partial t}[p,q]=\sum_{l=1,3}\left(\frac{\partial^2x_l}{\partial p\,\partial t}\frac{\partial\dot{x}_l}{\partial q}+\frac{\partial x_l}{\partial p}\frac{\partial^2\dot{x}_l}{\partial q\,\partial t}-\frac{\partial^2x_l}{\partial q\,\partial t}\frac{\partial\dot{x}_l}{\partial p}-\frac{\partial x_l}{\partial q}\frac{\partial^2\dot{x}_l}{\partial p\,\partial t}\right), \tag{11.43}$$

or

$$\frac{\partial}{\partial t}[p,q]=\sum_{l=1,3}\left[\frac{\partial}{\partial p}\left(\frac{\partial x_l}{\partial t}\frac{\partial\dot{x}_l}{\partial q}-\frac{\partial x_l}{\partial q}\frac{\partial\dot{x}_l}{\partial t}\right)-\frac{\partial}{\partial q}\left(\frac{\partial x_l}{\partial t}\frac{\partial\dot{x}_l}{\partial p}-\frac{\partial x_l}{\partial p}\frac{\partial\dot{x}_l}{\partial t}\right)\right]. \tag{11.44}$$

However, in the above expression, x_l and $\dot{x}_l$ stand for coordinates and velocities of Keplerian orbits calculated with $c_1,\cdots,c_6$ treated as constants. Thus, we can write $\partial x_l/\partial t\equiv\dot{x}_l$ and $\partial\dot{x}/\partial t\equiv\ddot{x}_l$, giving

$$\frac{\partial}{\partial t}[p,q]=\sum_{l=1,3}\left[\frac{\partial}{\partial p}\left(\frac{1}{2}\frac{\partial\dot{x}_l^2}{\partial q}-\frac{\partial F_0}{\partial x_l}\frac{\partial x_l}{\partial q}\right)-\frac{\partial}{\partial q}\left(\frac{1}{2}\frac{\partial\dot{x}_l^2}{\partial p}-\frac{\partial F_0}{\partial x_l}\frac{\partial x_l}{\partial p}\right)\right], \tag{11.45}$$

since

$$\ddot{x}_l = \frac{\partial F_0}{\partial x_l}, \tag{11.46}$$

where $F_0 = \mu/r$. Expression (11.45) reduces to

$$\frac{\partial}{\partial t}[p, q] = \frac{1}{2}\frac{\partial^2 v^2}{\partial p\,\partial q} - \frac{\partial^2 F_0}{\partial p\,\partial q} - \frac{1}{2}\frac{\partial^2 v^2}{\partial q\,\partial p} + \frac{\partial^2 F_0}{\partial q\,\partial p} = 0, \tag{11.47}$$

where $v^2 = \sum_{l=1,3} \dot{x}_l^2$. Hence, we conclude that Largrange brackets are functions of the osculating orbital elements, $c_1, \ldots, c_6$, but are not explicit functions of t. It follows that we can evaluate these brackets at any convenient point in the orbit.

11.5 Transformation of Lagrange Brackets

Let our six orbital elements be the major radius, a, the eccentricity, e, the mean longitude at epoch, $\bar{\lambda}_0$, the inclination, I, the longitude of the ascending node, Ω, and the longitude of the perihelion, ϖ—see Section 4.10. The mean orbital angular velocity is $n = (\mu/a^3)^{1/2}$ [see Equation (4.96)]. Consider how a particular Lagrange bracket transforms under a rotation of the coordinate system (X, Y, Z), which was introduced in Section 4.10, about the Z-axis (looking up the axis). We can write

$$[p, q] = \frac{\partial(X, \dot{X})}{\partial(p, q)} + \frac{\partial(Y, \dot{Y})}{\partial(p, q)} + \frac{\partial(Z, \dot{Z})}{\partial(p, q)}, \tag{11.48}$$

where

$$\frac{\partial(a, b)}{\partial(c, d)} \equiv \frac{\partial a}{\partial c}\frac{\partial b}{\partial d} - \frac{\partial a}{\partial d}\frac{\partial b}{\partial c}. \tag{11.49}$$

Let the new coordinate system be (x', y', z'). A rotation about the Z-axis though an angle Ω brings the ascending node to the x'-axis—see Figure 4.4. The relation between the old and new coordinates is

$$X = x' \cos\Omega - y' \sin\Omega, \tag{11.50}$$

$$Y = x' \sin\Omega + y' \cos\Omega, \tag{11.51}$$

$$Z = z'. \tag{11.52}$$

The partial derivatives with respect to p can be written

$$\frac{\partial X}{\partial p} = A_1 \cos\Omega - B_1 \sin\Omega, \tag{11.53}$$

$$\frac{\partial Y}{\partial p} = B_1 \cos\Omega + A_1 \sin\Omega, \tag{11.54}$$

$$\frac{\partial \dot{X}}{\partial p} = C_1 \cos\Omega - D_1 \sin\Omega, \tag{11.55}$$

$$\frac{\partial \dot{Y}}{\partial p} = D_1 \cos\Omega + C_1 \sin\Omega, \tag{11.56}$$

where

$$A_1 = \frac{\partial x'}{\partial p} - y'\,\frac{\partial \Omega}{\partial p}, \tag{11.57}$$

$$B_1 = \frac{\partial y'}{\partial p} + x'\,\frac{\partial \Omega}{\partial p}, \tag{11.58}$$

$$C_1 = \frac{\partial \dot{x}'}{\partial p} - \dot{y}'\,\frac{\partial \Omega}{\partial p}, \tag{11.59}$$

$$D_1 = \frac{\partial \dot{y}'}{\partial p} + \dot{x}'\,\frac{\partial \Omega}{\partial p}. \tag{11.60}$$

Let A_2, B_2, C_2, and D_2 be the equivalent quantities obtained by replacing p by q in the above equations. It thus follows that

$$\frac{\partial(X,\dot{X})}{\partial(p,q)} = (A_1\,C_2 - A_2\,C_1)\,\cos^2\Omega + (B_1\,D_2 - B_2\,D_1)\,\sin^2\Omega$$
$$+(-A_1\,D_2 - B_1\,C_2 + A_2\,D_1 + B_2\,C_1)\,\sin\Omega\,\cos\Omega, \tag{11.61}$$

$$\frac{\partial(Y,\dot{Y})}{\partial(p,q)} = (B_1\,D_2 - B_2\,D_1)\,\cos^2\Omega + (A_1\,C_2 - A_2\,C_1)\,\sin^2\Omega$$
$$+(A_1\,D_2 + B_1\,C_2 - A_2\,D_1 - B_2\,C_1)\,\sin\Omega\,\cos\Omega. \tag{11.62}$$

Hence,

$$[p,q] = A_1\,C_2 - A_2\,C_1 + B_1\,D_2 - B_2\,D_1 + \frac{\partial(Z,\dot{Z})}{\partial(p,q)}. \tag{11.63}$$

Now,

$$\begin{aligned} A_1\,C_2 - A_2\,C_1 &= \left(\frac{\partial x'}{\partial p} - y'\,\frac{\partial\Omega}{\partial p}\right)\left(\frac{\partial \dot{x}'}{\partial q} - \dot{y}'\,\frac{\partial\Omega}{\partial q}\right) \\ &\quad - \left(\frac{\partial x'}{\partial q} - y'\,\frac{\partial\Omega}{\partial q}\right)\left(\frac{\partial \dot{x}'}{\partial p} - \dot{y}'\,\frac{\partial\Omega}{\partial p}\right) \\ &= \frac{\partial(x',\dot{x}')}{\partial(p,q)} + \left(-y'\,\frac{\partial \dot{x}'}{\partial q} + \dot{y}'\,\frac{\partial x'}{\partial q}\right)\frac{\partial\Omega}{\partial p} + \left(-\dot{y}'\,\frac{\partial x'}{\partial p} + y'\,\frac{\partial \dot{x}'}{\partial p}\right)\frac{\partial\Omega}{\partial q}. \end{aligned} \tag{11.64}$$

Similarly,

$$\begin{aligned} B_1\,D_2 - B_2\,D_1 &= \left(\frac{\partial y'}{\partial p} + x'\,\frac{\partial\Omega}{\partial p}\right)\left(\frac{\partial \dot{y}'}{\partial q} + \dot{x}'\,\frac{\partial\Omega}{\partial q}\right) \\ &\quad - \left(\frac{\partial y'}{\partial q} + x'\,\frac{\partial\Omega}{\partial q}\right)\left(\frac{\partial \dot{y}'}{\partial p} + \dot{x}'\,\frac{\partial\Omega}{\partial p}\right) \\ &= \frac{\partial(y',\dot{y}')}{\partial(p,q)} + \left(x'\,\frac{\partial \dot{y}'}{\partial q} - \dot{x}'\,\frac{\partial y'}{\partial q}\right)\frac{\partial\Omega}{\partial p} + \left(\dot{x}'\,\frac{\partial y'}{\partial p} - x'\,\frac{\partial \dot{y}'}{\partial p}\right)\frac{\partial\Omega}{\partial q}. \end{aligned} \tag{11.65}$$

Let

$$[p,q]' = \frac{\partial(x',\dot{x}')}{\partial(p,q)} + \frac{\partial(y',\dot{y}')}{\partial(p,q)} + \frac{\partial(z',\dot{z}')}{\partial(p,q)}. \tag{11.66}$$

Since $Z = z'$ and $\dot{Z} = \dot{z}'$, it follows that

$$\begin{aligned}
[p,q] &= [p,q]' + \left(x'\,\frac{\partial\dot{y}'}{\partial q} + \dot{y}'\,\frac{\partial x'}{\partial q} - y'\,\frac{\partial\dot{x}'}{\partial q} - \dot{x}'\,\frac{\partial y'}{\partial q}\right)\frac{\partial\Omega}{\partial p}\\
&\quad - \left(x'\,\frac{\partial\dot{y}'}{\partial p} + \dot{y}'\,\frac{\partial x'}{\partial p} - y'\,\frac{\partial\dot{x}'}{\partial p} - \dot{x}'\,\frac{\partial y'}{\partial p}\right)\frac{\partial\Omega}{\partial q}\\
&= [p,q]' + \frac{\partial(\Omega,x'\,\dot{y}' - y'\,\dot{x}')}{\partial(p,q)}.
\end{aligned} \tag{11.67}$$

However,

$$x'\,\dot{y}' - y'\,\dot{x}' = h\,\cos I = [\mu\,a\,(1-e^2)]^{1/2}\,\cos I \equiv \mathcal{G}, \tag{11.68}$$

because the left-hand side is the component of the angular momentum along the z'-axis. Of course, this axis is inclined at an angle I to the z-axis, which is parallel to the angular momentum vector. Thus, we obtain

$$[p,q] = [p,q]' + \frac{\partial(\Omega,\mathcal{G})}{\partial(p,q)}. \tag{11.69}$$

Consider a rotation of the coordinate system about the x'-axis. Let the new coordinate system be (x'', y'', z''). A rotation through an angle I brings the orbit into the x''-y'' plane—see Figure 4.4. By analogy with the previous analysis,

$$[p,q]' = [p,q]'' + \frac{\partial(I,y''\,\dot{z}'' - z''\,\dot{y}'')}{\partial(p,q)}. \tag{11.70}$$

However, z'' and $\dot{z}''$ are both zero, since the orbit lies in the x''-y'' plane. Hence,

$$[p,q]' = [p,q]''. \tag{11.71}$$

Consider, finally, a rotation of the coordinate system about the z''-axis. Let the final coordinate system be (x, y, z). A rotation through an angle $\varpi - \Omega$ brings the perihelion to the x-axis—see Figure 4.4. By analogy with the previous analysis,

$$[p,q]'' = [p,q]''' + \frac{\partial(\varpi-\Omega,x\,\dot{y} - y\,\dot{x})}{\partial(p,q)}. \tag{11.72}$$

However,

$$x\,\dot{y} - y\,\dot{x} = h = [\mu\,a\,(1-e^2)]^{1/2} \equiv H, \tag{11.73}$$

so, from (11.69) and (11.71),

$$[p,q] = [p,q]''' + \frac{\partial(\varpi-\Omega,H)}{\partial(p,q)} + \frac{\partial(\Omega,\mathcal{G})}{\partial(p,q)}. \tag{11.74}$$

It thus remains to calculate

$$[p, q]''' = \frac{\partial x}{\partial p}\frac{\partial \dot{x}}{\partial q} - \frac{\partial x}{\partial q}\frac{\partial \dot{x}}{\partial p} + \frac{\partial y}{\partial p}\frac{\partial \dot{y}}{\partial q} - \frac{\partial y}{\partial q}\frac{\partial \dot{y}}{\partial p}. \tag{11.75}$$

The coordinates $x = r\cos\theta$ and $y = r\sin\theta$, where r is radial distance from the Sun, and θ the true anomaly, are functions of the major radius, a, the ellipticity, e, and the mean anomaly $\mathcal{M} = \bar{\lambda}_0 - \varpi + n\,t$. Since the Lagrange brackets are independent of time, it is sufficient to evaluate them at $\mathcal{M} = 0$: *i.e.*, at the perihelion point. Now, it is easily demonstrated from Equations (4.69)–(4.71) that

$$x = a\,(1-e) + \mathcal{O}(\mathcal{M}^2), \tag{11.76}$$

$$y = a\,\mathcal{M}\left(\frac{1+e}{1-e}\right)^{1/2} + \mathcal{O}(\mathcal{M}^3), \tag{11.77}$$

$$\dot{x} = -a\,n\,\frac{\mathcal{M}}{(1-e)^2} + \mathcal{O}(\mathcal{M}^3), \tag{11.78}$$

$$\dot{y} = a\,n\left(\frac{1+e}{1-e}\right)^{1/2} + \mathcal{O}(\mathcal{M}^2) \tag{11.79}$$

at small $\mathcal{M}$. Hence, at $\mathcal{M} = 0$,

$$\frac{\partial x}{\partial a} = 1-e, \tag{11.80}$$

$$\frac{\partial x}{\partial e} = -a, \tag{11.81}$$

$$\frac{\partial y}{\partial(\bar{\lambda}_0 - \varpi)} = a\left(\frac{1+e}{1-e}\right)^{1/2}, \tag{11.82}$$

$$\frac{\partial \dot{x}}{\partial(\bar{\lambda}_0 - \varpi)} = -\frac{a\,n}{(1-e)^2}, \tag{11.83}$$

$$\frac{\partial \dot{y}}{\partial a} = -\frac{n}{2}\left(\frac{1+e}{1-e}\right)^{1/2}, \tag{11.84}$$

$$\frac{\partial \dot{y}}{\partial e} = a\,n\,(1+e)^{-1/2}\,(1-e)^{-3/2}, \tag{11.85}$$

because $n \propto a^{-3/2}$. All other partial derivatives are zero. Now, since in the $(x,\,y,\,z)$ coordinate system the orbit only depends on the elements a, e, and $\bar{\lambda}_0 - \varpi$, we can write

$$\begin{aligned} [p, q]''' &= \frac{\partial(a, e)}{\partial(p, q)}\left[\frac{\partial(x, \dot{x})}{\partial(a, e)} + \frac{\partial(y, \dot{y})}{\partial(a, e)}\right] \\ &\quad + \frac{\partial(e, \bar{\lambda}_0 - \varpi)}{\partial(p, q)}\left[\frac{\partial(x, \dot{x})}{\partial(e, \bar{\lambda}_0 - \varpi)} + \frac{\partial(y, \dot{y})}{\partial(e, \bar{\lambda}_0 - \varpi)}\right] \\ &\quad + \frac{\partial(\bar{\lambda}_0 - \varpi, a)}{\partial(p, q)}\left[\frac{\partial(x, \dot{x})}{\partial(\bar{\lambda}_0 - \varpi, a)} + \frac{\partial(y, \dot{y})}{\partial(\bar{\lambda}_0 - \varpi, a)}\right]. \end{aligned} \tag{11.86}$$

Substitution of the values of the derivatives evaluated at $\mathcal{M} = 0$ into the above expression yields

$$\frac{\partial(x,\dot{x})}{\partial(a,e)} + \frac{\partial(y,\dot{y})}{\partial(a,e)} = 0, \tag{11.87}$$

$$\frac{\partial(x,\dot{x})}{\partial(e,\bar{\lambda}_0-\varpi)} + \frac{\partial(y,\dot{y})}{\partial(e,\bar{\lambda}_0-\varpi)} = 0, \tag{11.88}$$

$$\frac{\partial(x,\dot{x})}{\partial(\bar{\lambda}_0-\varpi,a)} + \frac{\partial(y,\dot{y})}{\partial(\bar{\lambda}_0-\varpi,a)} = \frac{a\,n}{2}. \tag{11.89}$$

Thus,

$$[p,q]''' = \frac{\partial(\bar{\lambda}_0-\varpi,a)}{\partial(p,q)}\,\frac{n\,a}{2} = \frac{\partial(\bar{\lambda}_0-\varpi,a)}{\partial(p,q)}\,\frac{\mu^{1/2}}{2\,a^{1/2}} = \frac{\partial(\bar{\lambda}_0-\varpi,L)}{\partial(p,q)}, \tag{11.90}$$

where $L = (\mu\,a)^{1/2}$. Hence, from (11.74), we obtain

$$[p,q] = \frac{\partial(\bar{\lambda}_0-\varpi,L)}{\partial(p,q)} + \frac{\partial(\varpi-\Omega,H)}{\partial(p,q)} + \frac{\partial(\Omega,\mathcal{G})}{\partial(p,q)}. \tag{11.91}$$

11.6 Lagrange's Planetary Equations

Now,

$$L = (\mu\,a)^{1/2}, \tag{11.92}$$

$$H = [\mu\,a\,(1-e^2)]^{1/2}, \tag{11.93}$$

$$\mathcal{G} = [\mu\,a\,(1-e^2)]^{1/2}\,\cos I, \tag{11.94}$$

and $n\,a = (\mu/a)^{1/2}$. Hence,

$$\frac{\partial L}{\partial a} = \frac{n\,a}{2}, \tag{11.95}$$

$$\frac{\partial H}{\partial a} = \frac{n\,a}{2}\,(1-e^2)^{1/2}, \tag{11.96}$$

$$\frac{\partial H}{\partial e} = -n\,a^2\,e\,(1-e^2)^{-1/2}, \tag{11.97}$$

$$\frac{\partial \mathcal{G}}{\partial a} = \frac{n\,a}{2}\,(1-e^2)^{1/2}\,\cos I, \tag{11.98}$$

$$\frac{\partial \mathcal{G}}{\partial e} = -n\,a^2\,e\,(1-e^2)^{-1/2}\,\cos I, \tag{11.99}$$

$$\frac{\partial \mathcal{G}}{\partial I} = -n\,a^2\,(1-e^2)^{1/2}\,\sin I, \tag{11.100}$$

with all other partial derivatives zero. Thus, from (11.91), the only non-zero Lagrange brackets are

$$[\bar{\lambda}_0, a] = -[a, \bar{\lambda}_0] = \frac{n\,a}{2}, \tag{11.101}$$

$$[\varpi, a] = -[a, \varpi] = -\frac{n\,a}{2}\,[1 - (1-e^2)^{1/2}], \tag{11.102}$$

$$[\Omega, a] = -[a, \Omega] = -\frac{n\,a}{2}\,(1-e^2)^{1/2}\,(1-\cos I), \tag{11.103}$$

$$[\varpi, e] = -[e, \varpi] = -n\,a^2\,e\,(1-e^2)^{-1/2}, \tag{11.104}$$

$$[\Omega, e] = -[e, \Omega] = n\,a^2\,e\,(1-e^2)^{-1/2}\,(1-\cos I), \tag{11.105}$$

$$[\Omega, I] = -[I, \Omega] = -n\,a^2\,(1-e^2)^{1/2}\,\sin I. \tag{11.106}$$

Hence, Equations (11.39) yield

$$[a, \bar{\lambda}_0]\,\frac{d\bar{\lambda}_0}{dt} + [a, \varpi]\,\frac{d\varpi}{dt} + [a, \Omega]\,\frac{d\Omega}{dt} = \frac{\partial \mathcal{R}}{\partial a}, \tag{11.107}$$

$$[e, \varpi]\,\frac{d\varpi}{dt} + [e, \Omega]\,\frac{d\Omega}{dt} = \frac{\partial \mathcal{R}}{\partial e}, \tag{11.108}$$

$$[\bar{\lambda}_0, a]\,\frac{da}{dt} = \frac{\partial \mathcal{R}}{\partial \bar{\lambda}_0}, \tag{11.109}$$

$$[I, \Omega]\,\frac{d\Omega}{dt} = \frac{\partial \mathcal{R}}{\partial I}, \tag{11.110}$$

$$[\Omega, a]\,\frac{da}{dt} + [\Omega, e]\,\frac{de}{dt} + [\Omega, I]\,\frac{dI}{dt} = \frac{\partial \mathcal{R}}{\partial \Omega}, \tag{11.111}$$

$$[\varpi, a]\,\frac{da}{dt} + [\varpi, e]\,\frac{de}{dt} = \frac{\partial \mathcal{R}}{\partial \varpi}. \tag{11.112}$$

Finally, Equations (11.101)–(11.112) can be rearranged to give

$$\frac{da}{dt} = \frac{2}{n\,a}\,\frac{\partial \mathcal{R}}{\partial \bar{\lambda}_0}, \tag{11.113}$$

$$\frac{de}{dt} = -\frac{(1-e^2)^{1/2}}{n\,a^2\,e}\,[1-(1-e^2)^{1/2}]\,\frac{\partial \mathcal{R}}{\partial \bar{\lambda}_0} - \frac{(1-e^2)^{1/2}}{n\,a^2\,e}\,\frac{\partial \mathcal{R}}{\partial \varpi}, \tag{11.114}$$

$$\frac{d\bar{\lambda}_0}{dt} = -\frac{2}{n\,a}\,\frac{\partial \mathcal{R}}{\partial a} + \frac{(1-e^2)^{1/2}\,[1-(1-e^2)^{1/2}]}{n\,a^2\,e}\,\frac{\partial \mathcal{R}}{\partial e} + \frac{\tan(I/2)}{n\,a^2\,(1-e^2)^{1/2}}\,\frac{\partial \mathcal{R}}{\partial I}, \tag{11.115}$$

$$\frac{dI}{dt} = -\frac{\tan(I/2)}{n\,a^2\,(1-e^2)^{1/2}}\left(\frac{\partial \mathcal{R}}{\partial \bar{\lambda}_0} + \frac{\partial \mathcal{R}}{\partial \varpi}\right) - \frac{(1-e^2)^{-1/2}}{n\,a^2\,\sin I}\,\frac{\partial \mathcal{R}}{\partial \Omega}, \tag{11.116}$$

$$\frac{d\Omega}{dt} = \frac{1}{n\,a^2\,(1-e^2)^{1/2}\,\sin I}\,\frac{\partial \mathcal{R}}{\partial I}, \tag{11.117}$$

$$\frac{d\varpi}{dt} = \frac{(1-e^2)^{1/2}}{n\,a^2\,e}\,\frac{\partial\mathcal{R}}{\partial e} + \frac{\tan(I/2)}{n\,a^2\,(1-e^2)^{1/2}}\,\frac{\partial\mathcal{R}}{\partial I}. \tag{11.118}$$

Equations (11.113)–(11.118), which specify the time evolution of the osculating orbital elements of the first planet due to the disturbing influence of the second, are known collectively as *Lagrange's planetary equations*. Obviously, there is an analogous set of equations, involving primed quantities, that specify the time evolution of the osculating elements of the second planet due to the influence of the first.

11.7 Transformation of Lagrange's Equations

It can be seen, from Equations (11.113)–(11.118), that in the limit of small eccentricity, e, and small inclination, I, certain terms on the right-hand sides of Lagrange's equations become singular. This problem can be alleviated by defining the alternative orbital elements,

$$h = e\,\sin\varpi, \tag{11.119}$$

$$k = e\,\cos\varpi, \tag{11.120}$$

$$p = \sin I\,\sin\Omega, \tag{11.121}$$

$$q = \sin I\,\cos\Omega. \tag{11.122}$$

If we write Lagrange's equations in terms of these new elements, we obtain

$$\frac{da}{dt} = \frac{2}{n\,a}\,\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0}, \tag{11.123}$$

$$\frac{d\bar{\lambda}_0}{dt} = -\frac{2}{n\,a}\,\frac{\partial\mathcal{R}}{\partial a} + \frac{(1-e^2)^{1/2}}{n\,a^2\,[1+(1-e^2)^{1/2}]}\left(h\,\frac{\partial\mathcal{R}}{\partial h} + k\,\frac{\partial\mathcal{R}}{\partial k}\right) + \frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\left(p\,\frac{\partial\mathcal{R}}{\partial p} + q\,\frac{\partial\mathcal{R}}{\partial q}\right), \tag{11.124}$$

$$\frac{dh}{dt} = -\frac{(1-e^2)^{1/2}}{n\,a^2\,[1+(1-e^2)^{1/2}]}\,h\,\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0} + \frac{(1-e^2)^{1/2}}{n\,a^2}\,\frac{\partial\mathcal{R}}{\partial k} + \frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,k\left(p\,\frac{\partial\mathcal{R}}{\partial p} + q\,\frac{\partial\mathcal{R}}{\partial q}\right), \tag{11.125}$$

$$\frac{dk}{dt} = -\frac{(1-e^2)^{1/2}}{n\,a^2\,[1+(1-e^2)^{1/2}]}\,k\,\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0} - \frac{(1-e^2)^{1/2}}{n\,a^2}\,\frac{\partial\mathcal{R}}{\partial h} - \frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,h\left(p\,\frac{\partial\mathcal{R}}{\partial p} + q\,\frac{\partial\mathcal{R}}{\partial q}\right), \tag{11.126}$$

$$\frac{dp}{dt} = -\frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,p\left(\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0} + k\,\frac{\partial\mathcal{R}}{\partial h} - h\,\frac{\partial\mathcal{R}}{\partial k}\right)$$

$$+\frac{\cos I}{n\,a^2\,(1-e^2)^{1/2}}\,\frac{\partial\mathcal{R}}{\partial q}, \tag{11.127}$$

$$\begin{aligned}\frac{dq}{dt} &= -\frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,q\left(\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0}+k\,\frac{\partial\mathcal{R}}{\partial h}-h\,\frac{\partial\mathcal{R}}{\partial k}\right)\\ &\quad-\frac{\cos I}{n\,a^2\,(1-e^2)^{1/2}}\,\frac{\partial\mathcal{R}}{\partial p}.\end{aligned} \tag{11.128}$$

Note that the transformed equations now contain no singular terms in the limit $e, I \rightarrow 0$.

As we shall see in Section 11.9, the disturbing function takes the form $\mathcal{R}(a, h, k, p, q, \bar{\lambda})$, where $\bar{\lambda} = \bar{\lambda}_0 + n\,t$. It follows that

$$\frac{\partial\mathcal{R}}{\partial\bar{\lambda}_0} = \frac{\partial\mathcal{R}}{\partial\bar{\lambda}}, \tag{11.129}$$

$$\frac{\partial\mathcal{R}}{\partial a} = \frac{\partial\mathcal{R}}{\partial a}+\frac{\partial\mathcal{R}}{\partial\bar{\lambda}}\,\frac{dn}{da}\,t. \tag{11.130}$$

The explicit factor t appearing in the above equation is problematic. Fortunately, it can easily be eliminated by replacing the variable $\bar{\lambda}_0$ by $\bar{\lambda}$. In this scheme, Equations (11.123) and (11.124) transform to

$$\frac{da}{dt} = \frac{2}{n\,a}\,\frac{\partial\mathcal{R}}{\partial\bar{\lambda}}, \tag{11.131}$$

$$\begin{aligned}\frac{d\bar{\lambda}}{dt} &= n-\frac{2}{n\,a}\,\frac{\partial\mathcal{R}}{\partial a}+\frac{(1-e^2)^{1/2}}{n\,a^2\,[1+(1-e^2)^{1/2}]}\left(h\,\frac{\partial\mathcal{R}}{\partial h}+k\,\frac{\partial\mathcal{R}}{\partial k}\right)\\ &\quad+\frac{\cos I}{2\,n\,a^2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\left(p\,\frac{\partial\mathcal{R}}{\partial p}+q\,\frac{\partial\mathcal{R}}{\partial q}\right),\end{aligned} \tag{11.132}$$

respectively, where $\partial/\partial\bar{\lambda}$ is taken at constant a, and $\partial/\partial a$ at constant $\bar{\lambda}$.

11.8 Expansion of Lagrange's Equations

According to expression (11.8), the first planet's disturbing function can be written in the form

$$\mathcal{R} = \frac{\tilde{\mu}'}{a'}\,\mathcal{S}, \tag{11.133}$$

where $\mathcal{S}$ is $\mathcal{O}(1)$. Thus, since $\mu = n^2\,a^3$, Lagrange's equations for the first planet reduce to

$$\frac{d\ln a}{dt} = 2\,n\,\epsilon'\alpha\,\frac{\partial\mathcal{S}}{\partial\bar{\lambda}}, \tag{11.134}$$

$$\frac{d\bar{\lambda}}{dt} = n-2\,n\,\epsilon'\alpha^2\,\frac{\partial\mathcal{S}}{\partial\alpha}+\frac{n\,\epsilon'\alpha\,(1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]}\left(h\,\frac{\partial\mathcal{S}}{\partial h}+k\,\frac{\partial\mathcal{S}}{\partial k}\right)$$

$$+\frac{n\,\epsilon'\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\left(p\,\frac{\partial\mathcal{S}}{\partial p}+q\,\frac{\partial\mathcal{S}}{\partial q}\right), \tag{11.135}$$

$$\frac{dh}{dt} = -\frac{n\,\epsilon'\alpha\,(1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]}\,h\,\frac{\partial\mathcal{S}}{\partial\bar{\lambda}}+n\,\epsilon'\alpha\,(1-e^2)^{1/2}\,\frac{\partial\mathcal{S}}{\partial k}$$
$$+\frac{n\,\epsilon'\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,k\left(p\,\frac{\partial\mathcal{S}}{\partial p}+q\,\frac{\partial\mathcal{S}}{\partial q}\right), \tag{11.136}$$

$$\frac{dk}{dt} = -\frac{n\,\epsilon'\alpha\,(1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]}\,k\,\frac{\partial\mathcal{S}}{\partial\bar{\lambda}}-n\,\epsilon'\alpha\,(1-e^2)^{1/2}\,\frac{\partial\mathcal{S}}{\partial h}$$
$$-\frac{n\,\epsilon'\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,h\left(p\,\frac{\partial\mathcal{S}}{\partial p}+q\,\frac{\partial\mathcal{S}}{\partial q}\right), \tag{11.137}$$

$$\frac{dp}{dt} = -\frac{n\,\epsilon'\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,p\left(\frac{\partial\mathcal{S}}{\partial\bar{\lambda}}+k\,\frac{\partial\mathcal{S}}{\partial h}-h\,\frac{\partial\mathcal{S}}{\partial k}\right)$$
$$+\frac{n\,\epsilon'\alpha\,\cos I}{(1-e^2)^{1/2}}\,\frac{\partial\mathcal{S}}{\partial q}, \tag{11.138}$$

$$\frac{dq}{dt} = -\frac{n\,\epsilon'\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\,q\left(\frac{\partial\mathcal{S}}{\partial\bar{\lambda}}+k\,\frac{\partial\mathcal{S}}{\partial h}-h\,\frac{\partial\mathcal{S}}{\partial k}\right)$$
$$-\frac{n\,\epsilon'\alpha\,\cos I}{(1-e^2)^{1/2}}\,\frac{\partial\mathcal{S}}{\partial p}. \tag{11.139}$$

Here,

$$\epsilon' = \frac{m'}{M+m}, \tag{11.140}$$

$$\alpha = \frac{a}{a'}. \tag{11.141}$$

Now, the Sun is much more massive than any planet in the Solar System. It follows that the parameter ϵ' is very small compared to unity. Expansion of Equations (11.134)–(11.139) to *first-order* in ϵ' yields

$$\bar{\lambda}(t) = \bar{\lambda}_0 + n\,t + \bar{\lambda}^{(1)}(t), \tag{11.142}$$

where $\lambda^{(1)} \sim \mathcal{O}(\epsilon')$, and

$$\frac{da}{dt} = \epsilon'\,n\left[2\,\alpha\,\frac{\partial\mathcal{S}}{\partial\bar{\lambda}^{(0)}}\right], \tag{11.143}$$

$$\frac{d\bar{\lambda}^{(1)}}{dt} = \epsilon'\,n\left[-\frac{3}{2}\,a-2\,\alpha^2\,\frac{\partial\mathcal{S}}{\partial\alpha}+\frac{\alpha\,(1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]}\left(h\,\frac{\partial\mathcal{S}}{\partial h}+k\,\frac{\partial\mathcal{S}}{\partial k}\right)\right.$$
$$\left.+\frac{\alpha\,\cos I}{2\,\cos^2(I/2)\,(1-e^2)^{1/2}}\left(p\,\frac{\partial\mathcal{S}}{\partial p}+q\,\frac{\partial\mathcal{S}}{\partial q}\right)\right], \tag{11.144}$$

$$\frac{dh}{dt} = \epsilon' n \left[-\frac{\alpha (1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]} h \frac{\partial \mathcal{S}}{\partial \bar{\lambda}^{(0)}} + \alpha (1-e^2)^{1/2} \frac{\partial \mathcal{S}}{\partial k} + \frac{\alpha \cos I}{2 \cos^2(I/2) (1-e^2)^{1/2}} k \left(p \frac{\partial \mathcal{S}}{\partial p} + q \frac{\partial \mathcal{S}}{\partial q} \right) \right], \tag{11.145}$$

$$\frac{dk}{dt} = \epsilon' n \left[-\frac{\alpha (1-e^2)^{1/2}}{[1+(1-e^2)^{1/2}]} k \frac{\partial \mathcal{S}}{\partial \bar{\lambda}^{(0)}} - \alpha (1-e^2)^{1/2} \frac{\partial \mathcal{S}}{\partial h} - \frac{\alpha \cos I}{2 \cos^2(I/2) (1-e^2)^{1/2}} h \left(p \frac{\partial \mathcal{S}}{\partial p} + q \frac{\partial \mathcal{S}}{\partial q} \right) \right], \tag{11.146}$$

$$\frac{dp}{dt} = \epsilon' n \left[-\frac{\alpha \cos I}{2 \cos^2(I/2) (1-e^2)^{1/2}} p \left(\frac{\partial \mathcal{S}}{\partial \bar{\lambda}^{(0)}} + k \frac{\partial \mathcal{S}}{\partial h} - h \frac{\partial \mathcal{S}}{\partial k} \right) + \frac{\alpha \cos I}{(1-e^2)^{1/2}} \frac{\partial \mathcal{S}}{\partial q} \right], \tag{11.147}$$

$$\frac{dq}{dt} = \epsilon' n \left[-\frac{\alpha \cos I}{2 \cos^2(I/2) (1-e^2)^{1/2}} q \left(\frac{\partial \mathcal{S}}{\partial \bar{\lambda}^{(0)}} + k \frac{\partial \mathcal{S}}{\partial h} - h \frac{\partial \mathcal{S}}{\partial k} \right) - \frac{\alpha \cos I}{(1-e^2)^{1/2}} \frac{\partial \mathcal{S}}{\partial p} \right], \tag{11.148}$$

where $\bar{\lambda}^{(0)} = \bar{\lambda}_0 + n\,t$. In the following, for ease of notation, $\bar{\lambda}^{(0)}$ is written simply as $\bar{\lambda}$.

It can be seen from Table 4.1 that the Planets all possess orbits whose eccentricities, e, and inclinations, I (in radians), are small compared to unity, but large compared to the ratio of any planetary mass to that of the Sun. It follows that

$$1 \gg e,\ I \gg \epsilon', \tag{11.149}$$

which is our fundamental ordering of small quantities. Assuming that $I \sim \mathcal{O}(e)$, we can perform a secondary expansion in the small parameter e. It turns out that when the disturbing function is expanded to *second-order* in e it takes the general form (see Section 11.9)

$$\mathcal{S} = \mathcal{S}_0(\alpha, \bar{\lambda}) + \mathcal{S}_1(\alpha, \bar{\lambda}, h, k) + \mathcal{S}_2(\alpha, \bar{\lambda}, h, k, p, q). \tag{11.150}$$

Here, $\mathcal{S}_n$ is $\mathcal{O}(e^n)$. So, if we expand the right-hand sides of Equations (11.143)–(11.148) to *first-order* in e then we obtain

$$\frac{da}{dt} = \epsilon' n \left[2\alpha \frac{\partial(\mathcal{S}_0 + \mathcal{S}_1)}{\partial \bar{\lambda}} \right], \tag{11.151}$$

$$\frac{d\bar{\lambda}^{(1)}}{dt} = \epsilon' n \left[-\frac{3}{2} a - 2\alpha^2 \frac{\partial(\mathcal{S}_0 + \mathcal{S}_1)}{\partial \alpha} + \alpha \left(h \frac{\partial \mathcal{S}_1}{\partial h} + k \frac{\partial \mathcal{S}_1}{\partial k} \right) \right], \tag{11.152}$$

$$\frac{dh}{dt} = \epsilon' n \left[-\alpha h \frac{\partial \mathcal{S}_0}{\partial \bar{\lambda}} + \alpha \frac{\partial(\mathcal{S}_1 + \mathcal{S}_2)}{\partial k} \right], \tag{11.153}$$

$$\frac{dk}{dt} = \epsilon' n \left[-\alpha\, k\, \frac{\partial \mathcal{S}_0}{\partial \bar{\lambda}} - \alpha\, \frac{\partial(\mathcal{S}_1 + \mathcal{S}_2)}{\partial h} \right], \tag{11.154}$$

$$\frac{dp}{dt} = \epsilon' n \left[-\frac{\alpha}{2}\, p\, \frac{\partial \mathcal{S}_0}{\partial \bar{\lambda}} + \alpha\, \frac{\partial \mathcal{S}_2}{\partial q} \right], \tag{11.155}$$

$$\frac{dq}{dt} = \epsilon' n \left[-\frac{\alpha}{2}\, q\, \frac{\partial \mathcal{S}_0}{\partial \bar{\lambda}} - \alpha\, \frac{\partial \mathcal{S}_2}{\partial p} \right]. \tag{11.156}$$

Note that h, k, p, q are $\mathcal{O}(e)$, whereas α and $\bar{\lambda}$ are $\mathcal{O}(1)$.

Analogous arguments reveal that Lagrange's equations for the orbital elements of the second planet take the form

$$\frac{da'}{dt} = \epsilon\, n' \left[2\, \alpha^{-1}\, \frac{\partial(\mathcal{S}_0' + \mathcal{S}_1')}{\partial \bar{\lambda}'} \right], \tag{11.157}$$

$$\frac{d\bar{\lambda}^{(1)\,\prime}}{dt} = \epsilon\, n' \left[-\frac{3}{2}\, a' + 2\, \frac{\partial(\mathcal{S}_0' + \mathcal{S}_1')}{\partial \alpha} + \alpha^{-1} \left(h'\, \frac{\partial \mathcal{S}_1'}{\partial h'} + k'\, \frac{\partial \mathcal{S}_1'}{\partial k'} \right) \right], \tag{11.158}$$

$$\frac{dh'}{dt} = \epsilon\, n' \left[-\alpha^{-1}\, h'\, \frac{\partial \mathcal{S}_0'}{\partial \bar{\lambda}'} + \alpha^{-1}\, \frac{\partial(\mathcal{S}_1' + \mathcal{S}_2')}{\partial k'} \right], \tag{11.159}$$

$$\frac{dk'}{dt} = \epsilon\, n' \left[-\alpha^{-1}\, k'\, \frac{\partial \mathcal{S}_0'}{\partial \bar{\lambda}'} - \alpha^{-1}\, \frac{\partial(\mathcal{S}_1' + \mathcal{S}_2')}{\partial h'} \right], \tag{11.160}$$

$$\frac{dp'}{dt} = \epsilon\, n' \left[-\frac{\alpha^{-1}}{2}\, p'\, \frac{\partial \mathcal{S}_0'}{\partial \bar{\lambda}'} + \alpha^{-1}\, \frac{\partial \mathcal{S}_2'}{\partial q'} \right], \tag{11.161}$$

$$\frac{dq'}{dt} = \epsilon\, n' \left[-\frac{\alpha^{-1}}{2}\, q'\, \frac{\partial \mathcal{S}_0'}{\partial \bar{\lambda}'} - \alpha^{-1}\, \frac{\partial \mathcal{S}_2'}{\partial p'} \right]. \tag{11.162}$$

Here, all primed quantities have analogous definitions to the corresponding unprimed quantities, and *vice versa*.

11.9 Expansion of Disturbing Function

We now need to evaluate the disturbing functions for both planets, and then expand them to second-order in e, I, e', and I'. In the following, it is assumed that I, e', $I' \sim \mathcal{O}(e)$.

The disturbing functions (11.8) and (11.9) can be written $\mathcal{R} = (\tilde{\mu}'/a')\,\mathcal{S}$ and $\mathcal{R}' = (\tilde{\mu}/a)\,\mathcal{S}'$, where

$$\mathcal{S} = \mathcal{S}_D + \alpha\, S_E, \tag{11.163}$$

$$\mathcal{S}' = \alpha\, \mathcal{S}_D + \alpha^{-1} \mathcal{S}_I. \tag{11.164}$$

Here,

$$\mathcal{S}_D = \frac{a'}{|\mathbf{r}' - \mathbf{r}|}, \tag{11.165}$$

and

$$\mathcal{S}_E = -\left(\frac{r}{a}\right)\left(\frac{a'}{r'}\right)^2 \cos\psi, \tag{11.166}$$

$$\mathcal{S}_I = -\left(\frac{r'}{a'}\right)\left(\frac{a}{r}\right)^2 \cos\psi. \tag{11.167}$$

In the above, ψ is the angle subtended between the directions of $\mathbf{r}$ and $\mathbf{r}'$.

Now

$$\mathcal{S}_D = a'\left[r'^2 - 2\,r'\,r\,\cos\psi + r^2\right]^{-1/2}. \tag{11.168}$$

Let

$$\zeta = \frac{r-a}{a}, \tag{11.169}$$

$$\zeta' = \frac{r'-a'}{a'}, \tag{11.170}$$

$$\delta = \cos\psi - \cos(\vartheta - \vartheta'), \tag{11.171}$$

where $\vartheta = \varpi + \theta$ and $\vartheta' = \varpi' + \theta''$. Here, θ and θ' are the true anomalies of the first and second planets, respectively. We expect ζ and ζ' to both be $\mathcal{O}(e)$ [see Equation (11.181)], and δ to be $\mathcal{O}(e^2)$ [see Equation (11.191)]. We can write

$$\mathcal{S}_D = (1+\zeta')^{-1}\left[1 - 2\,\tilde{\alpha}\,\cos(\vartheta-\vartheta') + \tilde{\alpha}^2 - 2\,\tilde{\alpha}\,\delta\right]^{-1/2}, \tag{11.172}$$

where

$$\tilde{\alpha} = \left(\frac{1+\zeta}{1+\zeta'}\right)\alpha. \tag{11.173}$$

Expanding in e, and only retaining terms up to $\mathcal{O}(e^2)$, we obtain

$$\mathcal{S}_D \simeq (1+\zeta')^{-1}\left[F + (\tilde{\alpha}-\alpha)\,D\,F + \frac{1}{2}\,(\tilde{\alpha}-\alpha)^2\,D^2\,F\right] + \delta\,\alpha\,F^3, \tag{11.174}$$

where $D \equiv \partial/\partial\alpha$, and

$$F(\alpha, \vartheta-\vartheta') = \frac{1}{[1 - 2\,\alpha\,\cos(\vartheta-\vartheta') + \alpha^2]^{1/2}}. \tag{11.175}$$

Hence,

$$\mathcal{S}_D \simeq \left[(1-\zeta'+\zeta'^2) + (\zeta-\zeta'-2\,\zeta\,\zeta'+2\,\zeta'^2)\,\alpha\,D + \frac{1}{2}(\zeta^2-2\,\zeta\,\zeta'+\zeta'^2)\,\alpha^2\,D^2\right]F$$
$$+\delta\,\alpha\,F^3. \tag{11.176}$$

Now, we can expand F and F^3 as Fourier series in $\vartheta - \vartheta'$:

$$F(\alpha, \vartheta - \vartheta') = \frac{1}{2} \sum_{j=-\infty,\infty} b_{1/2}^{(j)}(\alpha) \cos[j\,(\vartheta - \vartheta')], \tag{11.177}$$

$$F^3(\alpha, \vartheta - \vartheta') = \frac{1}{2} \sum_{j=-\infty,\infty} b_{3/2}^{(j)}(\alpha) \cos[j\,(\vartheta - \vartheta')], \tag{11.178}$$

where

$$b_s^{(j)}(\alpha) = \frac{1}{\pi} \int_0^{2\pi} \frac{\cos(j\,\psi)\,d\psi}{[1 - 2\,\alpha\,\cos\psi + \alpha^2]^s}. \tag{11.179}$$

Incidentally, the $b_s^{(j)}$ are known as *Laplace coefficients*. Thus,

$$\mathcal{S}_D \simeq \frac{1}{2} \sum_{j=-\infty,\infty} \left\{ \left[(1 - \zeta' + \zeta'^2) + (\zeta - \zeta' - 2\,\zeta\,\zeta' + 2\,\zeta'^2)\,\alpha\,D \right. \right. \tag{11.180}$$

$$\left. \left. + \frac{1}{2}(\zeta^2 - 2\,\zeta\,\zeta' + \zeta'^2)\,\alpha^2\,D^2 \right] b_{1/2}^{(j)}(\alpha) + \delta\,\alpha\,b_{3/2}^{(j)}(\alpha) \right\} \cos[j\,(\vartheta - \vartheta')],$$

where D now denotes $d/d\alpha$.

Equation (4.71) gives

$$\zeta \equiv \frac{r}{a} - 1 \simeq -e\,\cos\mathcal{M} + \frac{e^2}{2}\,(1 - \cos 2\mathcal{M})$$

$$\simeq -e\,\cos(\bar{\lambda} - \varpi) + \frac{e^2}{2}\left[1 - \cos(2\,\bar{\lambda} - 2\,\varpi)\right] \tag{11.181}$$

to $\mathcal{O}(e^2)$. Here, $\mathcal{M} = \bar{\lambda} - \varpi$ is the mean anomaly, and $\bar{\lambda} = \bar{\lambda}_0 + n\,t$ the mean ecliptic longitude. Obviously, there is an analogous equation for ζ'. Moreover, from Equation (4.70), we have

$$\sin\theta \simeq \sin\mathcal{M} + e\,\sin 2\mathcal{M} + \frac{e^2}{8}\,(9\,\sin 3\mathcal{M} - 7\,\sin\mathcal{M}), \tag{11.182}$$

$$\cos\theta \simeq \cos\mathcal{M} + e\,(\cos 2\mathcal{M} - 1) + \frac{e^2}{8}\,(9\,\cos 3\mathcal{M} - 9\,\cos\mathcal{M}). \tag{11.183}$$

Hence,

$$\cos(\omega + \theta) \equiv \cos\omega\,\cos\theta - \sin\omega\,\sin\theta$$

$$\simeq \cos(\omega + \mathcal{M}) + e\,[\cos(\omega + 2\mathcal{M}) - \cos\omega] \tag{11.184}$$

$$+ \frac{e^2}{8}\,[-8\,\cos(\omega + \mathcal{M}) - \cos(\omega - \mathcal{M}) + 9\,\cos(\omega + 3\mathcal{M})],$$

and

$$\begin{aligned}\sin(\omega+\theta) &\equiv \sin\omega\,\cos\theta+\cos\omega\,\sin\theta\\ &\simeq \sin(\omega+\mathcal{M})+e\,[\sin(\omega+2\mathcal{M})-\sin\omega]\\ &\quad+\frac{e^2}{8}\,[-8\,\sin(\omega+\mathcal{M})+\sin(\omega-\mathcal{M})+9\,\sin(\omega+3\mathcal{M})]\,.\end{aligned}\tag{11.185}$$

Thus, Equations (4.79)–(4.81) yield

$$\begin{aligned}\frac{X}{r} &\simeq \cos(\omega+\Omega+\mathcal{M})+e\,[\cos(\omega+\Omega+2\mathcal{M})-\cos(\omega+\Omega)]\\ &\quad+\frac{e^2}{8}\,[9\,\cos(\omega+\Omega+3\mathcal{M})-\cos(\omega+\Omega-\mathcal{M})-8\,\cos(\omega+\Omega+\mathcal{M})]\\ &\quad+s^2\,[\cos(\omega-\Omega+\mathcal{M})-\cos(\omega+\Omega+\mathcal{M})]\,, &(11.186)\\ \frac{Y}{r} &\simeq \sin(\omega+\Omega+\mathcal{M})+e\,[\sin(\omega+\Omega+2\mathcal{M})-\sin(\omega+\Omega)]\\ &\quad+\frac{e^2}{8}\,[9\,\sin(\omega+\Omega+3\mathcal{M})-\sin(\omega+\Omega-\mathcal{M})-8\,\sin(\omega+\Omega+\mathcal{M})]\\ &\quad-s^2\,[\sin(\omega-\Omega+\mathcal{M})+\sin(\omega+\Omega+\mathcal{M})]\,, &(11.187)\\ \frac{Z}{r} &\simeq 2\,s\,\sin(\omega+\mathcal{M})+2\,e\,s\,[\sin(\omega+2\mathcal{M})-\sin\omega)]\,, &(11.188)\end{aligned}$$

where $s\equiv\sin(I/2)$ is assumed to be $\mathcal{O}(e)$. Here, X, Y, Z are the Cartesian component of $\mathbf{r}$. There are, of course, completely analogous expressions for the Cartesian components of $\mathbf{r}'$.

Now,

$$\cos\psi=\frac{X}{r}\,\frac{X'}{r'}+\frac{Y}{r}\,\frac{Y'}{r'}+\frac{Z}{r}\,\frac{Z'}{r'},\tag{11.189}$$

so

$$\begin{aligned}\cos\psi &\simeq (1-e^2-e'^2-s^2-s'^2)\,\cos(\bar\lambda-\bar\lambda')+e\,e'\,\cos(2\,\bar\lambda-2\,\bar\lambda'-\varpi+\varpi')\\ &\quad+e\,e'\,\cos(\varpi-\varpi')+2\,s\,s'\,\cos(\bar\lambda-\bar\lambda'-\Omega+\Omega')\\ &\quad+e\,\cos(2\,\bar\lambda-\bar\lambda'-\varpi)-e\,\cos(\bar\lambda'-\varpi)+e'\,\cos(\bar\lambda-2\,\bar\lambda'+\varpi')\\ &\quad-e'\,\cos(\bar\lambda-\varpi')+\frac{9\,e^2}{8}\,\cos(3\,\bar\lambda-\bar\lambda'-2\,\varpi)-\frac{e^2}{8}\,\cos(\bar\lambda+\bar\lambda'-2\,\varpi)\\ &\quad+\frac{9\,e'^2}{8}\,\cos(\bar\lambda-3\,\bar\lambda'+2\,\varpi')-\frac{e'^2}{8}\,\cos(\bar\lambda+\bar\lambda'-2\,\varpi')\\ &\quad-e\,e'\,\cos(2\,\bar\lambda-\varpi-\varpi')-e\,e'\,\cos(2\,\bar\lambda'-\varpi-\varpi')\\ &\quad+s^2\,\cos(\bar\lambda+\bar\lambda'-2\,\Omega)+s'^2\,\cos(\bar\lambda+\bar\lambda'-2\,\Omega')\\ &\quad-2\,s\,s'\,\cos(\bar\lambda+\bar\lambda'-\Omega-\Omega').\end{aligned}\tag{11.190}$$

It is easily demonstrated that $\cos(\vartheta-\vartheta')$ represents the value taken by $\cos\psi$ when $s = s' = 0$. Hence, from (11.171) and (11.190),

$$\begin{aligned}\delta &\simeq s^2\left[\cos(\bar\lambda+\bar\lambda'-2\,\Omega)-\cos(\bar\lambda-\bar\lambda')\right]\\ &\quad +2\,s\,s'\left[\cos(\bar\lambda-\bar\lambda'-\Omega+\Omega')-\cos(\bar\lambda+\bar\lambda'-\Omega-\Omega')\right]\\ &\quad +s'^2\left[\cos(\bar\lambda+\bar\lambda'-2\,\Omega')-\cos(\bar\lambda-\bar\lambda')\right].\end{aligned}\tag{11.191}$$

Now,

$$\cos[j\,(\vartheta-\vartheta')]\equiv\cos(j\,\vartheta)\,\cos(j\,\vartheta')+\sin(j\,\vartheta)\,\sin(j\,\vartheta').\tag{11.192}$$

However, from (4.70),

$$\begin{aligned}\cos(j\,\vartheta) &\equiv \cos[j\,(\varpi+\theta)]\\ &\simeq (1-j^2\,e^2)\,\cos(j\,\bar\lambda)+e^2\left(\frac{j^2}{2}-\frac{5\,j}{8}\right)\cos[(2-j)\,\bar\lambda-2\,\varpi]\\ &\quad +e^2\left(\frac{j^2}{2}+\frac{5\,j}{8}\right)\cos[(2+j)\,\bar\lambda-2\,\varpi]\\ &\quad -j\,e\,\cos[(1-j)\,\bar\lambda-\varpi]+j\,e\,\cos[(1+j)\,\bar\lambda-\varpi)],\end{aligned}\tag{11.193}$$

since $\mathcal{M}=\bar\lambda-\varpi$. Likewise,

$$\begin{aligned}\sin(j\,\vartheta) &\equiv \sin[j\,(\varpi+\theta)]\\ &\simeq (1-j^2\,e^2)\,\sin(j\,\bar\lambda)+e^2\left(\frac{5\,j}{8}-\frac{j^2}{2}\right)\sin[(2-j)\,\bar\lambda-2\,\varpi]\\ &\quad +e^2\left(\frac{5\,j}{8}+\frac{j^2}{2}\right)\sin[(2+j)\,\bar\lambda-2\,\varpi]\\ &\quad +j\,e\,\sin[(1-j)\,\bar\lambda-\varpi]+j\,e\,\sin[(1+j)\,\bar\lambda-\varpi].\end{aligned}\tag{11.194}$$

Hence, we obtain

$$\begin{aligned}\cos[j\,(\vartheta-\vartheta')] &\simeq (1-j^2\,e^2-j^2\,e'^2)\,\cos[j\,(\bar\lambda-\bar\lambda')]\\ &\quad +e^2\left(\frac{5\,j}{8}+\frac{j^2}{2}\right)\cos[(2+j)\,\bar\lambda-j\,\bar\lambda'-2\,\varpi]\\ &\quad +e^2\left(\frac{j^2}{2}-\frac{5\,j}{8}\right)\cos[(2-j)\,\bar\lambda+j\,\bar\lambda'-2\,\varpi]\\ &\quad +j\,e\,\cos[(1+j)\,\bar\lambda-j\,\bar\lambda'-\varpi]-j\,e\,\cos[(1-j)\,\bar\lambda+j\,\bar\lambda'-\varpi]\\ &\quad +e'^2\left(\frac{j^2}{2}-\frac{5\,j}{8}\right)\cos[j\,\bar\lambda+(2-j)\,\bar\lambda'-2\,\varpi']\end{aligned}$$

$$+e'^2\left(\frac{5\,j}{8}+\frac{j^2}{2}\right)\cos[j\,\bar\lambda-(2+j)\,\bar\lambda'+2\,\varpi']$$
$$-j\,e'\,\cos[j\,\bar\lambda+(1-j)\,\bar\lambda'-\varpi']+j\,e'\,\cos[j\,\bar\lambda-(1+j)\,\bar\lambda'+\varpi']$$
$$-j^2\,e\,e'\,\cos[(1+j)\,\bar\lambda+(1-j)\,\bar\lambda'-\varpi-\varpi']$$
$$-j^2\,e\,e'\,\cos[(1-j)\,\bar\lambda+(1+j)\,\bar\lambda'-\varpi-\varpi']$$
$$+j^2\,e\,e'\,\cos[(1+j)\,\bar\lambda-(1+j)\,\bar\lambda'-\varpi+\varpi']$$
$$+j^2\,e\,e'\,\cos[(1-j)\,\bar\lambda-(1-j)\,\bar\lambda'-\varpi+\varpi']. \qquad (11.195)$$

Equations (11.180), (11.181), (11.191), and (11.195) yield

$$\mathcal{S}_D=\sum_{j=-\infty,\infty}\mathcal{S}^{(j)}, \qquad (11.196)$$

where

$$\mathcal{S}^{(j)} \simeq \left[\frac{b_{1/2}^{(j)}}{2}+\frac{1}{8}\left(e^2+e'^2\right)\left(-4\,j^2+2\,\alpha\,D+\alpha^2\,D^2\right)b_{1/2}^{(j)}\right.$$
$$\left.-\frac{\alpha}{4}\left(s^2+s'^2\right)\left(b_{3/2}^{(j-1)}+b_{3/2}^{(j+1)}\right)\right]\cos[j\,(\bar\lambda'-\bar\lambda)]$$
$$+\frac{e\,e'}{4}\left(2+6\,j+4\,j^2-2\,\alpha\,D-\alpha^2\,D^2\right)b_{1/2}^{(j+1)}\,\cos(j\,\bar\lambda'-j\,\bar\lambda+\varpi'-\varpi)$$
$$+s\,s'\,\alpha\,b_{3/2}^{(j+1)}\,\cos(j\,\bar\lambda'-j\,\bar\lambda+\Omega'-\Omega)$$
$$+\frac{e}{2}\,(-2\,j-\alpha\,D)\,b_{1/2}^{(j)}\,\cos[j\,\bar\lambda'+(1-j)\,\bar\lambda-\varpi]$$
$$+\frac{e'}{2}\,(-1+2\,j+\alpha\,D)\,b_{1/2}^{(j-1)}\,\cos[j\,\bar\lambda'+(1-j)\,\bar\lambda-\varpi']$$
$$+\frac{e^2}{8}\left(-5\,j+4\,j^2-2\,\alpha\,D+4\,j\,\alpha\,D+\alpha^2\,D^2\right)b_{1/2}^{(j)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-2\,\varpi]$$
$$+\frac{e\,e'}{4}\left(-2+6\,j-4\,j^2+2\,\alpha\,D-4\,j\,\alpha\,D-\alpha^2\,D^2\right)b_{1/2}^{(j-1)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-\varpi'-\varpi]$$
$$+\frac{e'^2}{8}\left(2-7\,j+4\,j^2-2\,\alpha\,D+4\,j\,\alpha\,D+\alpha^2\,D^2\right)b_{1/2}^{(j-2)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-2\,\varpi']$$
$$+\frac{s^2}{2}\,\alpha\,b_{3/2}^{(j-1)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-2\,\Omega]$$
$$-s\,s'\,\alpha\,b_{3/2}^{(j-1)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-\Omega'-\Omega]$$
$$+\frac{s'^2}{2}\,\alpha\,b_{3/2}^{(j-1)}\,\cos[j\,\bar\lambda'+(2-j)\,\bar\lambda-2\,\Omega']. \qquad (11.197)$$

Likewise, Equations (11.166), (11.167), (11.181), and (11.190) give

$$\mathcal{S}_E \simeq \left(-1+\frac{e^2}{2}+\frac{e'^2}{2}+s^2+s'^2\right)\cos(\bar\lambda'-\bar\lambda)$$

$$
\begin{aligned}
&-e\,e'\cos(2\,\bar\lambda'-2\,\bar\lambda-\varpi'+\varpi)-2\,s\,s'\cos(\bar\lambda'-\bar\lambda-\Omega'+\Omega)\\
&-\frac{e}{2}\cos(\bar\lambda'-2\,\bar\lambda+\varpi)+\frac{3\,e}{2}\cos(\bar\lambda'-\varpi)-2\,e'\cos(2\,\bar\lambda'-\bar\lambda-\varpi')\\
&-\frac{3\,e^2}{8}\cos(\bar\lambda'-3\,\bar\lambda+2\,\varpi)-\frac{e^2}{8}\cos(\bar\lambda'+\bar\lambda-2\,\varpi)\\
&+3\,e\,e'\cos(2\,\bar\lambda-\varpi'-\varpi)-\frac{e'^2}{8}\cos(\bar\lambda'+\bar\lambda-2\,\varpi')\\
&-\frac{27\,e'^2}{8}\cos(3\,\bar\lambda'-\bar\lambda-2\,\varpi')-s^2\cos(\bar\lambda'+\bar\lambda-2\,\Omega)\\
&+2\,s\,s'\cos(\bar\lambda'+\bar\lambda-\Omega'-\Omega)-s'^2\cos(\bar\lambda'+\bar\lambda-2\,\Omega'),
\end{aligned}
\tag{11.198}
$$

and

$$
\begin{aligned}
\mathcal{S}_I \simeq\;&\left(-1+\frac{e^2}{2}+\frac{e'^2}{2}+s^2+s'^2\right)\cos(\bar\lambda'-\bar\lambda)\\
&-e\,e'\cos(2\,\bar\lambda'-2\,\bar\lambda-\varpi'+\varpi)-2\,s\,s'\cos(\bar\lambda'-\bar\lambda-\Omega'+\Omega)\\
&-2\,e\cos(\bar\lambda'-2\,\bar\lambda+\varpi)+\frac{3\,e'}{2}\cos(\bar\lambda-\varpi')-\frac{e'}{2}\cos(2\,\bar\lambda'-\bar\lambda-\varpi')\\
&-\frac{27\,e^2}{8}\cos(\bar\lambda'-3\,\bar\lambda+2\,\varpi)-\frac{e^2}{8}\cos(\bar\lambda'+\bar\lambda-2\,\varpi)\\
&+3\,e\,e'\cos(2\,\bar\lambda-\varpi'-\varpi)-\frac{e'^2}{8}\cos(\bar\lambda'+\bar\lambda-2\,\varpi')\\
&-\frac{3\,e'^2}{8}\cos(3\,\bar\lambda'-\bar\lambda-2\,\varpi')-s^2\cos(\bar\lambda'+\bar\lambda-2\,\Omega)\\
&+2\,s\,s'\cos(\bar\lambda'+\bar\lambda-\Omega'-\Omega)-s'^2\cos(\bar\lambda'+\bar\lambda-2\,\Omega').
\end{aligned}
\tag{11.199}
$$

We can distinguish two different types of term that appear in our expansion of the disturbing function. *Periodic terms* vary *sinusoidally* in time as our two planets orbit the Sun (*i.e.*, they depend on the mean ecliptic longitudes, $\bar\lambda$ and $\bar\lambda'$), whereas *secular terms* remain *constant* in time (*i.e.*, they do not depend on $\bar\lambda$ and $\bar\lambda'$). We expect the periodic terms to give rise to relatively short-period (*i.e.*, of order the orbital periods) oscillations in the osculating orbital elements of our planets. On the other hand, we expect the secular terms to produce an initially linear increase in these elements with time. Since such an increase can become significant over a long period, even if the rate of increase is small, it is necessary to evaluate the secular terms in the disturbing function to higher order than the periodic terms. Hence, in the following, we shall evaluate periodic terms to $\mathcal{O}(e)$, and secular terms to $\mathcal{O}(e^2)$.

Making use of Equations (11.150), (11.163), and (11.196)–(11.198), as well as the definitions (11.119)–(11.122) (with $p \simeq I\sin\Omega$ and $q \simeq I\cos\omega$), the order by order

expansion (in e) of the disturbing function of the first planet is

$$\mathcal{S}_0 = \frac{1}{2}\sum_j b_{1/2}^{(j)}\cos[j\,(\bar\lambda-\bar\lambda')] - \alpha\,\cos(\bar\lambda-\bar\lambda'), \tag{11.200}$$

$$\begin{aligned}\mathcal{S}_1 = {}& \frac{1}{2}\sum_j \Big\{k\,(-2\,j+\alpha\,D)\,b_{1/2}^{(j)}\cos[(1-j)\,\bar\lambda+j\,\bar\lambda']\\ &+h\,(-2\,j+\alpha\,D)\,b_{1/2}^{(j)}\sin[(1-j)\,\bar\lambda+j\,\bar\lambda']\\ &+k'\,(-1+2\,j+\alpha\,D)\,b_{1/2}^{(j-1)}\cos[(1-j)\,\bar\lambda+j\,\bar\lambda']\\ &+h'\,(-1+2\,j+\alpha\,D)\,b_{1/2}^{(j-1)}\sin[(1-j)\,\bar\lambda+j\,\bar\lambda']\Big\}\\ &+\frac{\alpha}{2}\left\{-k\,\cos(2\,\bar\lambda-\bar\lambda')-h\,\sin(2\,\bar\lambda-\bar\lambda')+3\,k\,\cos\bar\lambda'+3\,h\,\sin\bar\lambda'\right.\\ &\left.-4\,k'\,\cos(\bar\lambda-2\,\bar\lambda')+4\,h'\,\sin(\bar\lambda-2\,\bar\lambda')\right\},\end{aligned} \tag{11.201}$$

$$\begin{aligned}\mathcal{S}_2 = {}& \frac{1}{8}\,(h^2+k^2+h'^2+k'^2)\,(2\,\alpha\,D+\alpha^2\,D^2)\,b_{1/2}^{(0)} - \frac{1}{8}\,(p^2+q^2+p'^2+q'^2)\,\alpha\,b_{3/2}^{(1)}\\ &+\frac{1}{4}\,(k\,k'+h\,h')\,(2-2\,\alpha\,D-\alpha^2\,D^2)\,b_{1/2}^{(1)}+\frac{1}{4}\,(p\,p'+q\,q')\,\alpha\,b_{3/2}^{(1)}.\end{aligned} \tag{11.202}$$

Likewise, the order by order expansion of the disturbing function of the second planet is

$$\mathcal{S}_0' = \frac{\alpha}{2}\sum_j b_{1/2}^{(j)}\cos[j\,(\bar\lambda'-\bar\lambda)] - \alpha^{-1}\,\cos(\bar\lambda'-\bar\lambda), \tag{11.203}$$

$$\begin{aligned}\mathcal{S}_1' = {}& \frac{\alpha}{2}\sum_j \Big\{k\,(-2\,j+\alpha\,D)\,b_{1/2}^{(j)}\cos[j\,\bar\lambda'+(1-j)\,\bar\lambda]\\ &+h\,(-2\,j+\alpha\,D)\,b_{1/2}^{(j)}\sin[j\,\bar\lambda'+(1-j)\,\bar\lambda]\\ &+k'\,(-1+2\,j+\alpha\,D)\,b_{1/2}^{(j-1)}\cos[j\,\bar\lambda'+(1-j)\,\bar\lambda]\\ &+h'\,(-1+2\,j+\alpha\,D)\,b_{1/2}^{(j-1)}\sin[j\,\bar\lambda'+(1-j)\,\bar\lambda]\Big\}\\ &+\frac{\alpha^{-1}}{2}\left\{-k'\,\cos(2\,\bar\lambda'-\bar\lambda)-h'\,\sin(2\,\bar\lambda'-\bar\lambda)+3\,k'\,\cos\bar\lambda+3\,h'\,\sin\bar\lambda\right.\\ &\left.-4\,k\,\cos(\bar\lambda'-2\,\bar\lambda)+4\,h\,\sin(\bar\lambda'-2\,\bar\lambda)\right\},\end{aligned} \tag{11.204}$$

$$\begin{aligned}\mathcal{S}_2' = {}& \frac{1}{8}\,(h^2+k^2+h'^2+k'^2)\,\alpha\,(2\,\alpha\,D+\alpha^2\,D^2)\,b_{1/2}^{(0)} - \frac{1}{8}\,(p^2+q^2+p'^2+q'^2)\,\alpha^2\,b_{3/2}^{(1)}\\ &+\frac{1}{4}\,(k\,k'+h\,h')\,\alpha\,(2-2\,\alpha\,D-\alpha^2\,D^2)\,b_{1/2}^{(1)}\\ &+\frac{1}{4}\,(p\,p'+q\,q')\,\alpha^2\,b_{3/2}^{(1)}.\end{aligned} \tag{11.205}$$

11.10 Secular Evolution of Planetary Orbits

As a specific example of the use of perturbation theory, let evaluate the long term evolution of the osculating orbital elements of the planets in our model Solar System due to the *secular* terms in their disturbing functions. This is equivalent to averaging the osculating elements over the relatively short time-scales associated with the periodic terms. From Equations (11.200)–(11.202), the secular part of the first planet's disturbing function takes the form

$$\begin{aligned}\mathcal{S}_s &= \frac{1}{2}\,b_{1/2}^{(0)} + \frac{1}{8}\,(h^2+k^2+h'^2+k'^2)\,\alpha\,b_{3/2}^{(1)} - \frac{1}{8}\,(p^2+q^2+p'^2+q'^2)\,\alpha\,b_{3/2}^{(1)}\\ &\quad -\frac{1}{4}\,(k\,k'+h\,h')\,\alpha\,b_{3/2}^{(2)} + \frac{1}{4}\,(p\,p'+q\,q')\,\alpha\,b_{3/2}^{(1)}, \qquad (11.206)\end{aligned}$$

since, as is easily demonstrated,

$$(2\,\alpha\,D + \alpha^2\,D^2)\,b_{1/2}^{(0)} = \alpha\,b_{3/2}^{(1)}, \qquad (11.207)$$

$$(2 - 2\,\alpha\,D - \alpha^2\,D^2)\,b_{1/2}^{(0)} = -\alpha\,b_{3/2}^{(2)}. \qquad (11.208)$$

Evaluating the right-hand sides of Equations (11.151)–(11.156) to $\mathcal{O}(e)$, we find that

$$\frac{da}{dt} = 0, \qquad (11.209)$$

$$\frac{d\bar{\lambda}^{(1)}}{dt} = \epsilon'\,n\left[-\frac{3}{2}\,a - 2\,\alpha^2\,\frac{\partial\mathcal{S}_s}{\partial\alpha}\right], \qquad (11.210)$$

$$\frac{dh}{dt} = \epsilon'\,n\left[\alpha\,\frac{\partial\mathcal{S}_s}{\partial k}\right], \qquad (11.211)$$

$$\frac{dk}{dt} = \epsilon'\,n\left[-\alpha\,\frac{\partial\mathcal{S}_s}{\partial h}\right], \qquad (11.212)$$

$$\frac{dp}{dt} = \epsilon'\,n\left[\alpha\,\frac{\partial\mathcal{S}_s}{\partial q}\right], \qquad (11.213)$$

$$\frac{dq}{dt} = \epsilon'\,n\left[-\alpha\,\frac{\partial\mathcal{S}_s}{\partial p}\right], \qquad (11.214)$$

since $\partial\mathcal{S}_s/\partial\bar{\lambda} = 0$. It follows that

$$a = a_0 - \epsilon'\,n\left[\frac{2}{3}\,\alpha^2\,D\,b_{1/2}^{(0)}\right], \qquad (11.215)$$

$$\bar{\lambda}_s^{(1)} = 0, \qquad (11.216)$$

$$\frac{dh}{dt} = \epsilon'\,n\left[\frac{1}{4}\,k\,\alpha^2\,b_{3/2}^{(1)} - \frac{1}{4}\,k'\,\alpha^2\,b_{3/2}^{(2)}\right], \qquad (11.217)$$

$$\frac{dk}{dt} = \epsilon' n \left[-\frac{1}{4} h\,\alpha^2 b_{3/2}^{(1)} + \frac{1}{4} h'\,\alpha^2 b_{3/2}^{(2)}\right], \tag{11.218}$$

$$\frac{dp}{dt} = \epsilon' n \left[-\frac{1}{4} q\,\alpha^2 b_{3/2}^{(1)} + \frac{1}{4} q'\,\alpha^2 b_{3/2}^{(1)}\right], \tag{11.219}$$

$$\frac{dq}{dt} = \epsilon' n \left[\frac{1}{4} p\,\alpha^2 b_{3/2}^{(1)} - \frac{1}{4} p'\,\alpha^2 b_{3/2}^{(1)}\right], \tag{11.220}$$

where a_0 is a constant. Here, we have assumed, without loss of generality, that the first planet's mean orbital angular velocity is not changed by the perturbation (*i.e.*, $d\bar{\lambda}_s^{(1)}/dt = 0$). Note that the first planet's major radius, a, remains constant in time.

From Equations (11.203)–(11.205), the secular part of the second planet's disturbing function takes the form

$$\begin{aligned}\mathcal{S}_s' &= \frac{1}{2}\,\alpha\, b_{1/2}^{(0)} + \frac{1}{8}\,(h^2+k^2+h'^2+k'^2)\,\alpha^2 b_{3/2}^{(1)} - \frac{1}{8}\,(p^2+q^2+p'^2+q'^2)\,\alpha^2 b_{3/2}^{(1)}\\ &\quad -\frac{1}{4}\,(k\,k'+h\,h')\,\alpha^2 b_{3/2}^{(2)} + \frac{1}{4}\,(p\,p'+q\,q')\,\alpha^2 b_{3/2}^{(1)}.\end{aligned} \tag{11.221}$$

Evaluating the right-hand sides of Equations (11.157)–(11.162) to $\mathcal{O}(\epsilon)$, we find that

$$\frac{da'}{dt} = 0, \tag{11.222}$$

$$\frac{d\bar{\lambda}^{(1)\prime}}{dt} = \epsilon\, n' \left[-\frac{3}{2}\,a' + 2\,\frac{\partial \mathcal{S}_s'}{\partial \alpha}\right], \tag{11.223}$$

$$\frac{dh'}{dt} = \epsilon\, n' \left[\alpha^{-1}\,\frac{\partial \mathcal{S}_s'}{\partial k'}\right], \tag{11.224}$$

$$\frac{dk'}{dt} = \epsilon\, n' \left[-\alpha^{-1}\,\frac{\partial \mathcal{S}_s'}{\partial h'}\right], \tag{11.225}$$

$$\frac{dp'}{dt} = \epsilon\, n' \left[\alpha^{-1}\,\frac{\partial \mathcal{S}_s'}{\partial q'}\right], \tag{11.226}$$

$$\frac{dq'}{dt} = \epsilon\, n' \left[-\alpha^{-1}\,\frac{\partial \mathcal{S}_s'}{\partial p'}\right]. \tag{11.227}$$

It follows that

$$a' = a_0' + \epsilon\, n' \left[\frac{2}{3}\,D\,(\alpha\, b_{1/2}^{(0)})\right], \tag{11.228}$$

$$\bar{\lambda}^{(1)\prime} = 0, \tag{11.229}$$

$$\frac{dh'}{dt} = \epsilon\, n' \left[\frac{1}{4}\,k'\,\alpha\, b_{3/2}^{(1)} - \frac{1}{4}\,k\,\alpha\, b_{3/2}^{(2)}\right], \tag{11.230}$$

$$\frac{dk'}{dt} = \epsilon\, n' \left[-\frac{1}{4}\, h'\,\alpha\, b_{3/2}^{(1)} + \frac{1}{4}\, h\,\alpha\, b_{3/2}^{(2)}\right], \tag{11.231}$$

$$\frac{dp'}{dt} = \epsilon\, n' \left[-\frac{1}{4}\, q'\,\alpha\, b_{3/2}^{(1)} + \frac{1}{4}\, q\,\alpha\, b_{3/2}^{(1)}\right], \tag{11.232}$$

$$\frac{dq'}{dt} = \epsilon\, n' \left[\frac{1}{4}\, p'\,\alpha\, b_{3/2}^{(1)} - \frac{1}{4}\, p\,\alpha\, b_{3/2}^{(1)}\right], \tag{11.233}$$

where a_0' is a constant. Note that the second planet's major radius, a', also remains constant in time.

Let us now generalize the above analysis to take all of the major planets in the Solar System into account. Let planet i have mass m_i, major radius a_i, eccentricity e_i, longitude of the perihelion ϖ_i, orbital inclination I_i, and longitude of the ascending node Ω_i. As we have seen, it is helpful to define the alternative orbital elements $h_i = e_i \sin\varpi_i$, $k_i = e_i \cos\varpi_i$, $p_i = \sin I_i \sin\Omega_i$, and $q_i = \sin I_i \cos\Omega_i$. It is also helpful to define the following parameters:

$$\alpha_{ij} = \begin{cases} a_i/a_j & a_j > a_i \\ a_j/a_i & a_j < a_i \end{cases}, \tag{11.234}$$

and

$$\bar{\alpha}_{ij} = \begin{cases} a_i/a_j & a_j > a_i \\ 1 & a_j < a_i \end{cases}, \tag{11.235}$$

as well as

$$\epsilon_{ij} = \frac{m_j}{M + m_i}, \tag{11.236}$$

where M is the mass of the Sun. It then follows, from the above analysis, that the secular terms in the planetary disturbing functions cause the h_i, k_i, p_i, and q_i to vary in time as

$$\frac{dh_i}{dt} = \sum_{j=1,8} A_{ij}\, k_j, \tag{11.237}$$

$$\frac{dk_i}{dt} = -\sum_{j=1,8} A_{ij}\, h_j, \tag{11.238}$$

$$\frac{dp_i}{dt} = \sum_{j=1,8} B_{ij}\, q_j, \tag{11.239}$$

$$\frac{dq_i}{dt} = -\sum_{j=1,8} B_{ij}\, p_j, \tag{11.240}$$

where

$$A_{ii} = \sum_{j\neq i} \frac{n_i}{4}\, \epsilon_{ij}\, \alpha_{ij}\, \bar{\alpha}_{ij}\, b_{3/2}^{(1)}(\alpha_{ij}), \tag{11.241}$$

$$A_{ij} = -\frac{n_i}{4}\,\epsilon_{ij}\,\alpha_{ij}\,\bar{\alpha}_{ij}\,b_{3/2}^{(2)}(\alpha_{ij}), \tag{11.242}$$

$$B_{ii} = -\sum_{j\neq i}\frac{n_i}{4}\,\epsilon_{ij}\,\alpha_{ij}\,\bar{\alpha}_{ij}\,b_{3/2}^{(1)}(\alpha_{ij}), \tag{11.243}$$

$$B_{ij} = \frac{n_i}{4}\,\epsilon_{ij}\,\alpha_{ij}\,\bar{\alpha}_{ij}\,b_{3/2}^{(1)}(\alpha_{ij}), \tag{11.244}$$

for $j \neq i$. Here, Mercury is planet 1, Venus is planet 2, *etc.*, and Neptune is planet 8. Note that the time evolution of the h_i and the k_i, which determine the eccentricities of the planetary orbits, is decoupled from that of the p_i and the q_i, which determine the inclinations. Let us search for normal mode solutions to Equations (11.237)–(11.240) of the form

$$h_i = \sum_{l=1,8} e_{il}\,\sin(g_l\,t+\beta_l), \tag{11.245}$$

$$k_i = \sum_{l=1,8} e_{il}\,\cos(g_l\,t+\beta_l), \tag{11.246}$$

$$p_i = \sum_{l=1,8} I_{il}\,\sin(f_l\,t+\gamma_l), \tag{11.247}$$

$$q_i = \sum_{l=1,8} I_{il}\,\cos(f_l\,t+\gamma_l). \tag{11.248}$$

It follows that

$$\sum_{j=1,8}(A_{ij}-\delta_{ij}\,g_l)\,e_{jl} = 0, \tag{11.249}$$

$$\sum_{j=1,8}(B_{ij}-\delta_{ij}\,f_l)\,I_{jl} = 0. \tag{11.250}$$

At this stage, we have effectively reduced the problem of determining the secular evolution of the planetary orbits to a pair of matrix eigenvalue equations that can be solved via standard numerical methods. Once we have determined the eigenfrequencies, f_l and g_l, and the corresponding eigenvectors, e_{il} and I_{il}, the phase angles β_l and γ_l are found by demanding that, at $t = 0$, Equations (11.245)–(11.248) lead to the values of e_i, I_i, Ω_i, and ϖ_i given in Table 4.1.

The theory outlined above is generally referred to as *Laplace-Lagrange secular evolution theory*. The eigenfrequencies, eigenvectors, and phase angles obtained from this theory are shown in Tables 11.1–11.3. Note that the largest eigenfrequency is of magnitude 25.89 arc seconds per year, which translates to an oscillation period of about 5×10^4 years. In other words, the time-scale over which the secular evolution of the Solar System predicted by Laplace-Lagrange theory takes place is at least 5×10^4 years.

Note that one of the inclination eigenfrequencies, f_5, takes the value zero. This is a consequence of the conservation of angular momentum. Since the Solar System is effectively

l	$g_l('' y^{-1})$	$f_l('' y^{-1})$	$\beta_l(^\circ)$	$\gamma_l(^\circ)$
1	5.460	-5.199	89.64	20.23
2	7.343	-6.568	195.0	318.3
3	17.32	-18.74	336.2	255.6
4	18.00	-17.63	319.0	296.9
5	3.723	0.000	30.14	107.5
6	22.43	-25.89	131.1	127.3
7	2.709	-2.911	110.5	315.6
8	0.6336	-0.6780	67.19	202.8

Table 11.1: *Eigenfrequencies (in arc seconds per year) and phase angles obtained from Laplace-Lagrange secular evolution theory.*

an isolated dynamical system, its net angular momentum vector, **L**, is constant in both magnitude and direction. The plane normal to **L** that passes through the center of mass of the Solar System (which lies very close to the Sun) is known as the *invariable plane*. If all of the planetary orbits were to lie in the invariable plane then the net angular velocity vector of the Solar System would be parallel to the fixed net angular momentum vector. Moreover, the angular momentum vector would be parallel to one of the Solar System's principle axes of rotation. In this situation, we would expect the Solar System to remain in a steady state (see Chapter 8). In other words, we would not expect any time evolution of the planetary inclinations. (Of course, lack of time variation implies an eigenfrequency of zero.) According to Equations (11.247) and (11.248), and the data shown in Tables 11.1 and 11.3, if the Solar System were in the inclination eigenstate associated with the null eigenfrequency f_5 then we would have

$$p_i = 2.751 \times 10^{-2} \sin\gamma_5, \tag{11.251}$$

$$q_i = 2.751 \times 10^{-2} \cos\gamma_5, \tag{11.252}$$

for $i = 1, 8$. Since $p_i = \sin I_i \sin\Omega_i$ and $q_i = \sin I_i \cos\Omega_i$, it follows that all of the planetary orbits would lie in the same plane, and this plane—which is, of course, the invariable plane—is inclined at $I_5 = \sin^{-1}(2.751 \times 10^{-2}) = 1.576^\circ$ to the ecliptic plane. Furthermore, the longitude of the ascending node of the invariable plane, with respect to the ecliptic plane, is $\Omega_5 = \gamma_5 = 107.5^\circ$. Actually, it is generally more convenient to measure the inclinations of the planetary orbits with respect to the invariable plane, rather than the ecliptic plane, since the inclination of the latter plane varies in time. We can achieve this goal by simply omitting the fifth inclination eigenstate when calculating orbital inclinations from Equations (11.247) and (11.248).

Consider the ith planet. Suppose one of the e_{il} coefficients—e_{ik}, say—is sufficiently

	$i = 1$	2	3	4	5	6	7	8
$l = 1$	18131	629	404	65	0	0	0	0
2	-2328	1917	1496	265	-1	-1	0	0
3	154	-1262	1046	2978	0	0	0	0
4	-169	1489	-1485	7286	0	0	0	0
5	2445	1635	1634	1877	4330	3415	-4395	158
6	10	-51	243	1566	-1563	4840	-181	-13
7	58	57	60	80	202	185	2934	-314
8	0	1	1	2	6	6	143	943

Table 11.2: *Components of the eccentricity eigenvectors* e_{il} *obtained from Laplace-Lagrange secular evolution theory. All components have been multiplied by* 10^5.

large that

$$|e_{ik}| > \sum_{l=1,8}^{l \neq k} |e_{il}|. \tag{11.253}$$

This is known as the *Lagrange condition*. As is easily demonstrated, if the Lagrange condition is satisfied then the eccentricity of the ith planet's orbit varies between the minimum value

$$e_{i\,\rm min} = |e_{ik}| - \sum_{l=1,8}^{l \neq k} |e_{il}| \tag{11.254}$$

and the maximum value

$$e_{i\,\rm max} = |e_{ik}| + \sum_{l=1,8}^{l \neq k} |e_{il}|. \tag{11.255}$$

Moreover, on average, the perihelion point precesses at the associated eigenfrequency, g_k. The precession is prograde (*i.e.*, in the same direction as the orbital motion) if the frequency is positive, and retrograde (*i.e.*, in the opposite direction) if the frequency is negative. If the Lagrange condition is not satisfied then all we can say is that the maximum eccentricity is given by Equation (11.255), and there is no minimum eccentricity (*i.e.*, the eccentricity can vary all the way down to zero). Furthermore, no mean precession rate of the perihelion point can be identified. It can be seen from Table 11.2 that the Lagrange condition for the orbital eccentricities is satisfied for all planets except for Venus and the Earth. The maximum and minimum eccentricities, and mean perihelion precession rates, of the Planets (when they exist) are given in Table 11.4. Note that Jupiter and Uranus have the same mean perihelion precession rates, and that all planets that possess mean precession rates exhibit prograde precession.

There is also a Lagrange condition associated with the inclinations of the planetary orbits. This condition is satisfied for the ith planet if one of the I_{il}—I_{ik}, say—is sufficiently

	$i = 1$	2	3	4	5	6	7	8
$l = 1$	12549	1180	850	180	-2	-2	2	0
2	-3547	1007	811	180	-1	-1	0	0
3	409	-2684	2445	-3596	0	0	0	0
4	116	-685	451	5022	0	-1	0	0
5	2751	2751	2751	2751	2751	2751	2751	2751
6	27	14	279	954	-636	1587	-69	-7
7	-332	-191	-172	-125	-95	-77	1757	-205
8	-144	-132	-129	-123	-117	-112	109	1181

Table 11.3: *Components of the inclination eigenvectors* I_{il} *obtained from Laplace-Lagrange secular evolution theory. All components have been multiplied by* 10^5.

large that

$$|I_{ik}| > \sum_{l=1,8}^{l\neq k, l\neq 5} |I_{il}|. \tag{11.256}$$

Note that the fifth inclination eigenmode is omitted from the above summation because we are now measuring inclinations relative to the invariable plane. If the Lagrange condition is satisfied then the inclination of the ith planet's orbit with respect to the invariable plane varies between the minimum value

$$I_{i\,\min} = \sin^{-1}\left[|I_{ik}| - \sum_{l=1,8}^{l\neq k, l\neq 5} |I_{il}|\right] \tag{11.257}$$

and the maximum value

$$I_{i\,\max} = \sin^{-1}\left[|I_{ik}| + \sum_{l=1,8}^{l\neq k, l\neq 5} |I_{il}|\right]. \tag{11.258}$$

Moreover, on average, the ascending node precesses at the associated eigenfrequency, f_k. The precession is prograde (*i.e.*, in the same direction as the orbital motion) if the frequency is positive, and retrograde (*i.e.*, in the opposite direction) if the frequency is negative. If the Lagrange condition is not satisfied then all we can say is that the maximum inclination is given by Equation (11.258), and there is no minimum inclination (*i.e.*, the inclination can vary all the way down to zero). Furthermore, no mean precession rate of the ascending node can be identified. It can be seen from Table 11.3 that the Lagrange condition for the orbital inclinations is satisfied for all planets except for Venus, the Earth, and Mars. The maximum and minimum inclinations, and mean nodal precession rates, of the Planets (when they exist) are given in Table 11.4. It can be seen that the four outer planets, which possess most of the mass of the Solar System, all have orbits whose inclinations to the invariable plane remain small. On the other hand, the four relatively light inner planets have orbits whose inclinations to the invariable plane can become relatively

Planet	e_{max}	e_{min}	$\langle\dot{\varpi}\rangle$	$I_{max}(^\circ)$	$I_{min}(^\circ)$	$\langle\dot{\Omega}\rangle$
Mercury	0.233	0.130	g_1	9.86	4.57	f_1
Venus	0.00704	–	–	3.38	–	–
Earth	0.0637	–	–	2.95	–	–
Mars	0.141	0.0450	g_4	5.84	–	–
Jupiter	0.0610	0.0256	g_5	0.488	0.241	f_6
Saturn	0.0845	0.0123	g_6	1.02	0.797	f_6
Uranus	0.0765	0.0114	g_5	1.11	0.904	f_7
Neptune	0.0143	0.00456	g_8	0.799	0.555	f_8

Table 11.4: *Maximum and minimum eccentricities and inclinations of the planetary orbits, as well as mean perihelion and nodal precession rates, obtained from Laplace-Lagrange secular evolution theory. All inclinations are relative to the invariable plane.*

large. Note that Jupiter and Saturn have the same mean nodal precession rates, and that all planets that possess mean precession rates exhibit retrograde precession.

11.11 Hirayama Families

Let us now consider the perturbing influence of the Planets on the orbit of an asteroid. Since all asteroids have masses that are much smaller than those of the Planets, it is reasonable to suppose that the perturbing influence of the asteroid on the planetary orbits is negligible. Let the asteroid have the osculating orbital elements a, e, $\bar{\lambda}_0$, I, Ω, ϖ—see Section 4.10—and the alternative elements h, k, p, q—see Section 11.7. Thus, the mean orbital angular velocity of the asteroid is $n = (G\,M/a^3)^{1/2}$. Likewise, let the eight planets have the osculating orbital elements a_i, e_i, $\bar{\lambda}_{0\,i}$, I_i, Ω_i, ϖ_i, and the alternative elements h_i, k_i, p_i, q_i, for $i = 1, 8$. It is helpful to define the following parameters:

$$\alpha_i = \begin{cases} a/a_i & a_i > a \\ a_i/a & a_i < a \end{cases}, \tag{11.259}$$

and

$$\bar{\alpha}_i = \begin{cases} a/a_i & a_i > a \\ 1 & a_i < a \end{cases}, \tag{11.260}$$

as well as

$$\epsilon_i = \frac{m_i}{M}, \tag{11.261}$$

where M is the mass of the Sun.

By analogy with the analysis in the previous section, the secular terms in the disturbing function of the asteroid, generated by the perturbing influence of the Planets, cause the

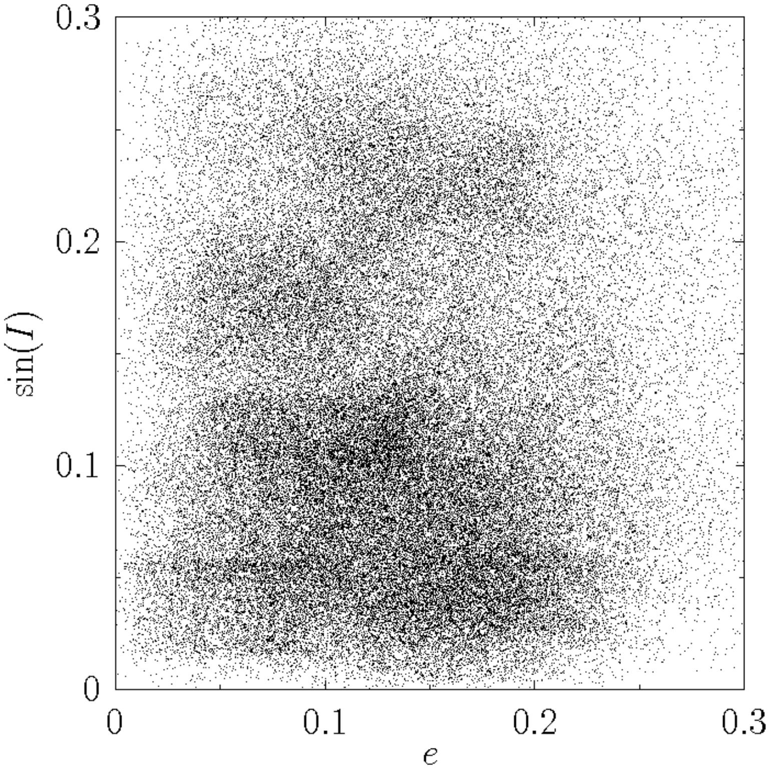


Figure 11.2: *Osculating eccentricity plotted against the sine of the osculating inclination (relative to the J2000 equinox and ecliptic) for the orbits of the first 100,000 numbered asteroids at MJD 55400. Data from JPL.*

asteroid's osculating orbital elements to evolve in time as

$$\frac{dh}{dt} = A\,k + \sum_{i=1,8} A_i\,k_i, \tag{11.262}$$

$$\frac{dk}{dt} = -A\,h - \sum_{i=1,8} A_i\,h_i, \tag{11.263}$$

$$\frac{dp}{dt} = B\,q + \sum_{i=1,8} B_i\,q_i, \tag{11.264}$$

$$\frac{dq}{dt} = -B\,p - \sum_{i=1,8} B_i\,p_i, \tag{11.265}$$

where

$$A = \sum_{i=1,8} \frac{n}{4}\,\epsilon_i\,\alpha_i\,\bar{\alpha}_i\,b_{3/2}^{(1)}(\alpha_i), \tag{11.266}$$

$$A_i = -\frac{n}{4}\,\epsilon_i\,\alpha_i\,\bar{\alpha}_i\,b_{3/2}^{(2)}(\alpha_i), \tag{11.267}$$

$$B = -\sum_{i=1,8} \frac{n}{4}\,\epsilon_i\,\alpha_i\,\bar{\alpha}_i\,b_{3/2}^{(1)}(\alpha_i), \tag{11.268}$$

$$B_i = \frac{n}{4}\,\epsilon_i\,\alpha_i\,\bar{\alpha}_i\,b_{3/2}^{(1)}(\alpha_i). \tag{11.269}$$

However, as we have seen, the planetary osculating elements themselves evolve in time as

$$h_i = \sum_{l=1,8} e_{il}\,\sin(g_l\,t + \beta_l), \tag{11.270}$$

$$k_i = \sum_{l=1,8} e_{il}\,\cos(g_l\,t + \beta_l), \tag{11.271}$$

$$p_i = \sum_{l=1,8} I_{il}\,\sin(f_l\,t + \gamma_l), \tag{11.272}$$

$$q_i = \sum_{l=1,8} I_{il}\,\cos(f_l\,t + \gamma_l). \tag{11.273}$$

Equations (11.262)–(11.265) can be solved to give

$$h(t) = e_{\rm free}\,\sin(A\,t + \beta_{\rm free}) + h_{\rm forced}(t), \tag{11.274}$$

$$k(t) = e_{\rm free}\,\cos(A\,t + \beta_{\rm free}) + k_{\rm forced}(t), \tag{11.275}$$

$$p(t) = \sin I_{\rm free}\,\sin(B\,t + \gamma_{\rm free}) + p_{\rm forced}(t), \tag{11.276}$$

$$q(t) = \sin I_{\rm free}\,\cos(B\,t + \gamma_{\rm free}) + q_{\rm forced}(t), \tag{11.277}$$

where

$$h_{\rm forced} = -\sum_{l=1,8} \frac{\nu_l}{A-g_l}\,\sin(g_l\,t+\beta_l), \tag{11.278}$$

$$k_{\rm forced} = -\sum_{l=1,8} \frac{\nu_l}{A-g_l}\,\cos(g_l\,t+\beta_l), \tag{11.279}$$

$$p_{\rm forced} = -\sum_{l=1,8} \frac{\mu_l}{B-f_l}\,\sin(f_l\,t+\gamma_l), \tag{11.280}$$

$$q_{\rm forced} = -\sum_{l=1,8} \frac{\mu_l}{B-f_l}\,\cos(f_l\,t+\gamma_l), \tag{11.281}$$

and

$$\nu_l = \sum_{i=1,8} A_i\,e_{il}, \tag{11.282}$$

$$\mu_l = \sum_{i=1,8} B_i\,I_{il}. \tag{11.283}$$

The parameters, $e_{\rm free}$ and $I_{\rm free}$, appearing in Equations (11.274)–(11.277), are the eccentricity and inclination, respectively, that the asteroid orbit would possess were it not for the perturbing influence of the Planets. Roughly speaking, the planetary perturbations cause the osculating eccentricity, $e = [h^2+k^2]^{1/2}$, and inclination, $I = \sin^{-1}([p^2+q^2]^{1/2})$, to oscillate about the intrinsic values $e_{\rm free}$ and $I_{\rm free}$, respectively.

Figure 11.2 shows the *osculating* eccentricity plotted against the sine of the *osculating* inclination for the orbits of the first 100,000 numbered asteroids (asteroids are numbered in order of their discovery). No particular patten is apparent. Figure 11.3 shows the *free* eccentricity plotted against the sine of the *free* inclination for the same 100,000 orbits. It can be seen that many of the points representing the asteroid orbits have condensed into clumps. These clumps are known as *Hirayama families*, after their discoverer, the Japanese astronomer Kiyotsugu Hirayama (1874-1943). It is thought that the asteroids making up a given family had a common origin—most likely due to the break up of some much larger body. As a consequence of this origin, the asteroids originally had similar orbital elements. However, as time progressed, these elements were jumbled by the perturbing influence of the Planets. Thus, it is only when this influence is removed that the commonality of the orbits becomes apparent. Hirayama families are named after their largest member. The most prominent families are the (4) Vesta, (15) Eunomia, (24) Themis, (44) Nysa, (158) Koronis, (221) Eos, and (1272) Gefion families. (The number in brackets is that of the corresponding asteroid.)

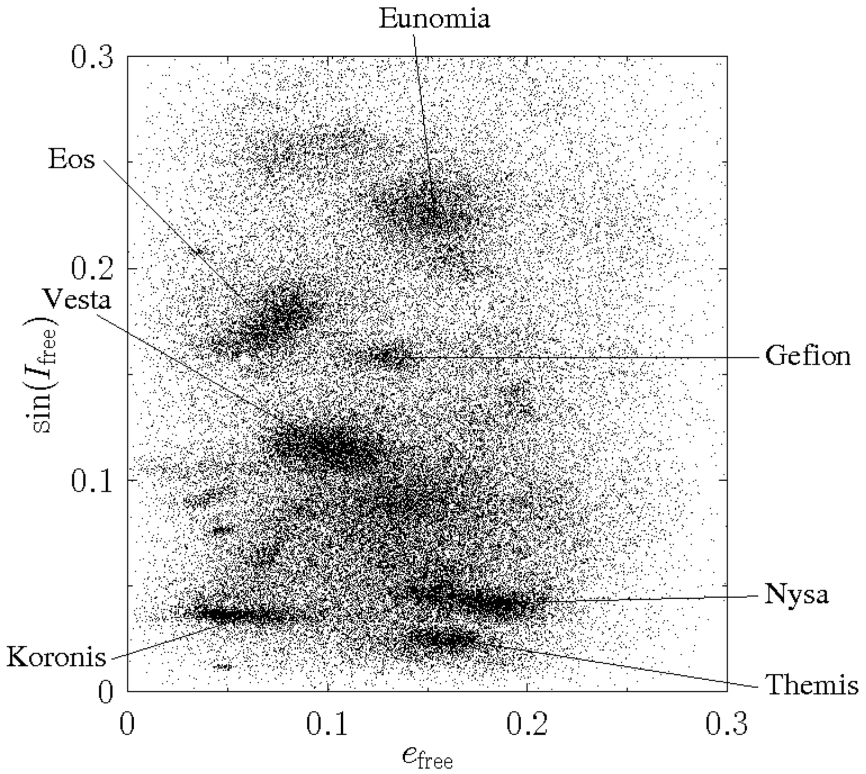


Figure 11.3: *Free eccentricity plotted against the sine of the free inclination (relative to the J2000 equinox and ecliptic) for the orbits of the first 100,000 numbered asteroids at MJD 55400. Data from JPL. The most prominent Hiroyama families are labeled.*

A Vector Algebra and Vector Calculus

A.1 Introduction

This appendix contains a brief outline of vector algebra and vector calculus. The essential purpose of *vector algebra* is to convert the propositions of Euclidean geometry in three-dimensional space into a convenient algebraic form. *Vector calculus* allows us to define the instantaneous velocity and acceleration of a moving object in three-dimensional space, as well as the work done when such an object travels along a general curved trajectory in a force field. Vector calculus also introduces the concept of a *scalar field*: *e.g.*, the potential energy associated with a conservative force field.

A.2 Scalars and Vectors

Many physical quantities (*e.g.*, mass, energy) are entirely defined by a numerical magnitude (in appropriate units). Such quantities, which have no directional element, are known as *scalars*. Moreover, since scalars can be represented by real numbers, it follows that they obey the familiar laws of ordinary algebra. However, there exits a second class of physical quantities (*e.g.*, velocity, acceleration, force) which are only completely defined when both a numerical magnitude *and* a direction in space is specified. Such quantities are known as *vectors*. By definition, a vector obeys the same algebra as a displacement in space, and may thus be represented geometrically by a straight-line, $\overrightarrow{PQ}$ (say), where the arrow indicates the direction of the displacement (*i.e.*, from point P to point Q)—see Figure A.1. The magnitude of the vector is represented by the length of the straight-line.

It is conventional to denote vectors by bold-faced symbols (*e.g.*, $\mathbf{a}$, $\mathbf{F}$) and scalars by non-bold-faced symbols (*e.g.*, r, S). The magnitude of a general vector, $\mathbf{a}$, is denoted $|\mathbf{a}|$, or just a, and is, by definition, always greater than or equal to zero. It is convenient to define a vector with zero magnitude—this is denoted $\mathbf{0}$, and has no direction. Finally, two vectors, $\mathbf{a}$ and $\mathbf{b}$, are said to be equal when their magnitudes and directions are identical.

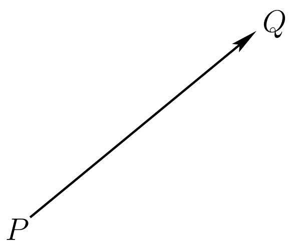


Figure A.1: *A vector.*

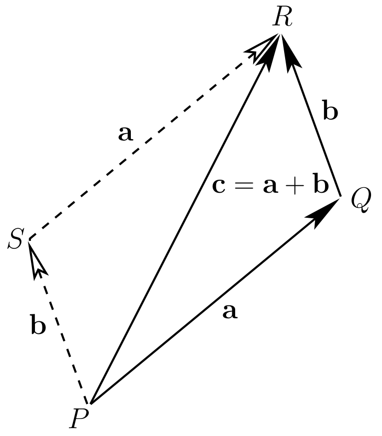


Figure A.2: *Vector addition.*

A.3 Vector Algebra

Suppose that the displacements $\overrightarrow{PQ}$ and $\overrightarrow{QR}$ represent the vectors **a** and **b**, respectively—see Figure A.2. It can be seen that the result of combining these two displacements is to give the net displacement $\overrightarrow{PR}$. Hence, if $\overrightarrow{PR}$ represents the vector **c** then we can write

$$\mathbf{c} = \mathbf{a} + \mathbf{b}. \tag{A.1}$$

This defines *vector addition*. By completing the parallelogram PQRS, we can also see that

$$\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PS} + \overrightarrow{SR}\,. \tag{A.2}$$

However, $\overrightarrow{PS}$ has the same length and direction as $\overrightarrow{QR}$, and, thus, represents the *same* vector, **b**. Likewise, $\overrightarrow{PQ}$ and $\overrightarrow{SR}$ both represent the vector **a**. Thus, the above equation is equivalent to

$$\mathbf{c} = \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \tag{A.3}$$

We conclude that the addition of vectors is *commutative*. It can also be shown that the *associative* law holds: *i.e.*,

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}. \tag{A.4}$$

The null vector, **0**, is represented by a displacement of zero length and arbitrary direction. Since the result of combining such a displacement with a finite length displacement is the same as the latter displacement by itself, it follows that

$$\mathbf{a} + \mathbf{0} = \mathbf{a}, \tag{A.5}$$

where **a** is a general vector. The negative of **a** is defined as that vector which has the same magnitude, but acts in the opposite direction, and is denoted $-\mathbf{a}$. The sum of **a** and $-\mathbf{a}$ is thus the null vector: *i.e.*,

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}. \tag{A.6}$$

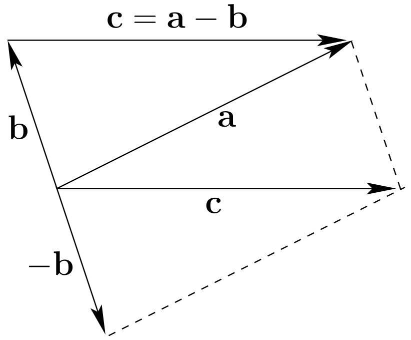


Figure A.3: *Vector subtraction.*

We can also define the difference of two vectors, **a** and **b**, as

$$\mathbf{c} = \mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}). \tag{A.7}$$

This definition of *vector subtraction* is illustrated in Figure A.3.

If $n > 0$ is a scalar then the expression $n\,\mathbf{a}$ denotes a vector whose direction is the same as **a**, and whose magnitude is n times that of **a**. (This definition becomes obvious when n is an integer.) If n is negative then, since $n\,\mathbf{a} = |n|\,(-\mathbf{a})$, it follows that $n\,\mathbf{a}$ is a vector whose magnitude is $|n|$ times that of **a**, and whose direction is opposite to **a**. These definitions imply that if n and m are two scalars then

$$n\,(m\,\mathbf{a}) = n\,m\,\mathbf{a} = m\,(n\,\mathbf{a}), \tag{A.8}$$

$$(n + m)\,\mathbf{a} = n\,\mathbf{a} + m\,\mathbf{a}, \tag{A.9}$$

$$n\,(\mathbf{a} + \mathbf{b}) = n\,\mathbf{a} + n\,\mathbf{b}. \tag{A.10}$$

A.4 Cartesian Components of a Vector

Consider a Cartesian coordinate system $Oxyz$ consisting of an origin, O, and three mutually perpendicular coordinate axes, Ox, Oy, and Oz—see Figure A.4. Such a system is said to be *right-handed* if, when looking along the Oz direction, a 90° clockwise rotation about Oz is required to take Ox into Oy. Otherwise, it is said to be left-handed. In physics, it is conventional to always use right-handed coordinate systems.

It is convenient to define unit vectors, $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, parallel to Ox, Oy, and Oz, respectively. Incidentally, a unit vector is a vector whose magnitude is unity. The position vector, **r**, of some general point P whose Cartesian coordinates are (x, y, z) is then given by

$$\mathbf{r} = x\,\mathbf{e}_z + y\,\mathbf{e}_y + z\,\mathbf{e}_z. \tag{A.11}$$

In other words, we can get from O to P by moving a distance x parallel to Ox, then a distance y parallel to Oy, and then a distance z parallel to Oz. Similarly, if **a** is an arbitrary vector then

$$\mathbf{a} = a_x\,\mathbf{e}_x + a_y\,\mathbf{e}_y + a_z\,\mathbf{e}_z, \tag{A.12}$$

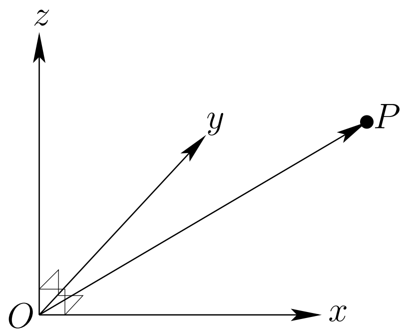


Figure A.4: *A right-handed Cartesian coordinate system.*

where a_x, a_y, and a_z are termed the *Cartesian components* of $\mathbf{a}$. It is coventional to write $\mathbf{a} \equiv (a_x,\, a_y,\, a_z)$. It follows that $\mathbf{e}_x \equiv (1,\, 0,\, 0)$, $\mathbf{e}_y \equiv (0,\, 1,\, 0)$, and $\mathbf{e}_z \equiv (0,\, 0,\, 1)$. Of course, $\mathbf{0} \equiv (0,\, 0,\, 0)$.

According to the three-dimensional generalization of the Pythagorean theorem, the distance $OP \equiv |\mathbf{r}| = r$ is given by

$$r = \sqrt{x^2 + y^2 + z^2}. \tag{A.13}$$

By analogy, the magnitude of a general vector $\mathbf{a}$ takes the form

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}. \tag{A.14}$$

If $\mathbf{a} \equiv (a_x,\, a_y,\, a_z)$ and $\mathbf{b} \equiv (b_x,\, b_y,\, b_z)$ then it is easily demonstrated that

$$\mathbf{a} + \mathbf{b} \equiv (a_x + b_x,\, a_y + b_y,\, a_z + b_z). \tag{A.15}$$

Furthermore, if n is a scalar then it is apparent that

$$n\,\mathbf{a} \equiv (n\,a_x,\, n\,a_y,\, n\,a_z). \tag{A.16}$$

A.5 Coordinate Transformations

A Cartesian coordinate system allows position and direction in space to be represented in a very convenient manner. Unfortunately, such a coordinate system also introduces arbitrary elements into our analysis. After all, two independent observers might well choose coordinate systems with different origins, and different orientations of the coordinate axes. In general, a given vector $\mathbf{a}$ will have different sets of components in these two coordinate systems. However, the direction and magnitude of $\mathbf{a}$ are the same in both cases. Hence, the two sets of components must be related to one another in a very particular fashion. Actually, since vectors are represented by *moveable* line elements in space (*i.e.*, in Figure A.2, $\overrightarrow{PQ}$ and $\overrightarrow{SR}$ represent the *same* vector), it follows that the components of a general vector

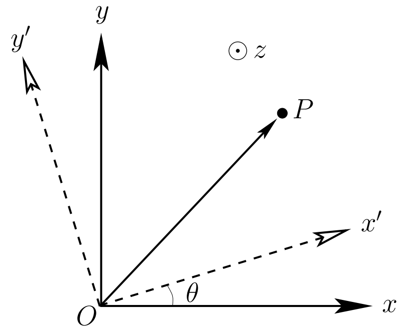


Figure A.5: *Rotation of the coordinate axes about* Oz.

are not affected by a simple shift in the origin of a Cartesian coordinate system. On the other hand, the components are modified when the coordinate axes are rotated.

Suppose that we transform to a new coordinate system, $Ox'y'z'$, which has the same origin as $Oxyz$, and is obtained by rotating the coordinate axes of $Oxyz$ through an angle θ about Oz—see Figure A.5. Let the coordinates of a general point P be (x, y, z) in $Oxyz$ and (x', y', z') in $Ox'y'z'$. According to simple trigonometry, these two sets of coordinates are related to one another via the transformation:

$$x' = x\cos\theta + y\sin\theta, \tag{A.17}$$

$$y' = -x\sin\theta + y\cos\theta, \tag{A.18}$$

$$z' = z. \tag{A.19}$$

Consider the vector displacement $\mathbf{r} \equiv \overrightarrow{OP}$. Note that this displacement is represented by the *same* symbol, **r**, in both coordinate systems, since the magnitude and direction of **r** are manifestly *independent* of the orientation of the coordinate axes. The coordinates of **r** *do* depend on the orientation of the axes: *i.e.*, $\mathbf{r} \equiv (x, y, z)$ in $Oxyz$, and $\mathbf{r} \equiv (x', y', z')$ in $Ox'y'z'$. However, they must depend in a very specific manner [*i.e.*, Equations (A.17)–(A.19)] which preserves the magnitude and direction of **r**.

The components of a general vector **a** transform in an analogous manner to Equations (A.17)–(A.19): *i.e.*,

$$a_{x'} = a_x\cos\theta + a_y\sin\theta, \tag{A.20}$$

$$a_{y'} = -a_x\sin\theta + a_y\cos\theta, \tag{A.21}$$

$$a_{z'} = a_z. \tag{A.22}$$

Moreover, there are similar transformation rules for rotation about Ox and Oy. Equations (A.20)–(A.22) effectively constitute the *definition* of a vector: *i.e.*, the three quantities (a_x, a_y, a_z) are the components of a vector provided that they transform under rotation of the coordinate axes about Oz in accordance with Equations (A.20)–(A.22). (And also transform correctly under rotation about Ox and Oy). Conversely, (a_x, a_y, a_z) *cannot* be

the components of a vector if they do not transform in accordance with Equations (A.20)–(A.22). Of course, scalar quantities are *invariant* under rotation of the coordinate axes. Thus, the individual components of a vector (a_x, say) are real numbers, but they are *not* scalars. Displacement vectors, and all vectors derived from displacements (*e.g.*, velocity, acceleration), automatically satisfy Equations (A.20)–(A.22). There are, however, other physical quantities which have both magnitude and direction, but which are not obviously related to displacements. We need to check carefully to see whether these quantities are really vectors (see Section A.8).

A.6 Scalar Product

A scalar quantity is invariant under all possible rotational transformations. The individual components of a vector are not scalars because they change under transformation. Can we form a scalar out of some combination of the components of one, or more, vectors? Suppose that we were to define the "percent" product,

$$\mathbf{a}\,\%\,\mathbf{b} \equiv a_x\, b_z + a_y\, b_x + a_z\, b_y = \text{scalar number}, \tag{A.23}$$

for general vectors $\mathbf{a}$ and $\mathbf{b}$. Is $\mathbf{a}\,\%\,\mathbf{b}$ invariant under transformation, as must be the case if it is a scalar number? Let us consider an example. Suppose that $\mathbf{a} \equiv (0,\, 1,\, 0)$ and $\mathbf{b} \equiv (1,\, 0,\, 0)$. It is easily seen that $\mathbf{a}\,\%\,\mathbf{b} = 1$. Let us now rotate the coordinate axes through 45° about Oz. In the new coordinate system, $\mathbf{a} \equiv (1/\sqrt{2},\, 1/\sqrt{2},\, 0)$ and $\mathbf{b} \equiv (1/\sqrt{2},\, -1/\sqrt{2},\, 0)$, giving $\mathbf{a}\,\%\,\mathbf{b} = 1/2$. Clearly, $\mathbf{a}\,\%\,\mathbf{b}$ is *not* invariant under rotational transformation, so the above definition is a bad one.

Consider, now, the *dot product* or *scalar product*:

$$\mathbf{a}\cdot\mathbf{b} \equiv a_x\, b_x + a_y\, b_y + a_z\, b_z = \text{scalar number}. \tag{A.24}$$

Let us rotate the coordinate axes though θ degrees about Oz. According to Equations (A.20)–(A.22), $\mathbf{a}\cdot\mathbf{b}$ takes the form

$$\begin{aligned}
\mathbf{a}\cdot\mathbf{b} &= (a_x \cos\theta + a_y \sin\theta)\,(b_x \cos\theta + b_y \sin\theta) \\
&\quad + (-a_x \sin\theta + a_y \cos\theta)\,(-b_x \sin\theta + b_y \cos\theta) + a_z\, b_z \\
&= a_x\, b_x + a_y\, b_y + a_z\, b_z
\end{aligned} \tag{A.25}$$

in the new coordinate system. Thus, $\mathbf{a}\cdot\mathbf{b}$ *is* invariant under rotation about Oz. It can easily be shown that it is also invariant under rotation about Ox and Oy. We conclude that $\mathbf{a}\cdot\mathbf{b}$ is a true scalar, and that the above definition is a good one. Incidentally, $\mathbf{a}\cdot\mathbf{b}$ is the *only* simple combination of the components of two vectors which transforms like a scalar. It is easily shown that the dot product is commutative and distributive: *i.e.*,

$$\begin{aligned}
\mathbf{a}\cdot\mathbf{b} &= \mathbf{b}\cdot\mathbf{a}, \\
\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) &= \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}.
\end{aligned} \tag{A.26}$$

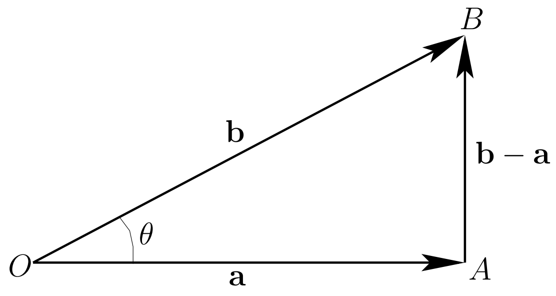


Figure A.6: *A vector triangle.*

The associative property is meaningless for the dot product, because we cannot have $(\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c}$, since $\mathbf{a}\cdot\mathbf{b}$ is scalar.

We have shown that the dot product $\mathbf{a}\cdot\mathbf{b}$ is coordinate independent. But what is the geometric significance of this? Well, in the special case where $\mathbf{a}=\mathbf{b}$, we get

$$\mathbf{a}\cdot\mathbf{b} = {a_x}^2 + {a_y}^2 + {a_z}^2 = |\mathbf{a}|^2 = a^2. \tag{A.27}$$

So, the invariance of $\mathbf{a}\cdot\mathbf{a}$ is equivalent to the invariance of the magnitude of vector $\mathbf{a}$ under transformation.

Let us now investigate the general case. The length squared of AB in the vector triangle shown in Figure A.6 is

$$(\mathbf{b}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{a}) = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\,\mathbf{a}\cdot\mathbf{b}. \tag{A.28}$$

However, according to the "cosine rule" of trigonometry,

$$(AB)^2 = (OA)^2 + (OB)^2 - 2\,(OA)\,(OB)\,\cos\theta, \tag{A.29}$$

where (AB) denotes the length of side AB. It follows that

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos\theta. \tag{A.30}$$

In this case, the invariance of $\mathbf{a}\cdot\mathbf{b}$ under transformation is equivalent to the invariance of the angle subtended between the two vectors. Note that if $\mathbf{a}\cdot\mathbf{b}=0$ then either $|\mathbf{a}|=0$, $|\mathbf{b}|=0$, or the vectors $\mathbf{a}$ and $\mathbf{b}$ are mutually perpendicular. The angle subtended between two vectors can easily be obtained from the dot product: *i.e.*,

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}. \tag{A.31}$$

The work W performed by a constant force $\mathbf{F}$ which moves an object through a displacement $\mathbf{r}$ is the product of the magnitude of $\mathbf{F}$ times the displacement in the direction of $\mathbf{F}$. If the angle subtended between $\mathbf{F}$ and $\mathbf{r}$ is θ then

$$W = |\mathbf{F}|\,(|\mathbf{r}|\,\cos\theta) = \mathbf{F}\cdot\mathbf{r}. \tag{A.32}$$

The work dW performed by a non-constant force $\mathbf{f}$ which moves an object through an infinitesimal displacement $d\mathbf{r}$ in a time interval dt is $dW = \mathbf{f}\cdot d\mathbf{r}$. Thus, the rate at which the force does work on the object, which is usually referred to as the power, is $P = dW/dt = \mathbf{f}\cdot d\mathbf{r}/dt$, or $P = \mathbf{f}\cdot\mathbf{v}$, where $\mathbf{v} = d\mathbf{r}/dt$ is the object's instantaneous velocity.

A.7 Vector Product

We have discovered how to construct a *scalar* from the components of two general vectors **a** and **b**. Can we also construct a *vector* which is not just a linear combination of **a** and **b**? Consider the following definition:

$$\mathbf{a} * \mathbf{b} \equiv (a_x\, b_x,\; a_y\, b_y,\; a_z\, b_z). \tag{A.33}$$

Is $\mathbf{a} * \mathbf{b}$ a proper vector? Suppose that $\mathbf{a} = (0,\, 1,\, 0)$, $\mathbf{b} = (1,\, 0,\, 0)$. In this case, $\mathbf{a} * \mathbf{b} = \mathbf{0}$. However, if we rotate the coordinate axes through 45° about Oz then $\mathbf{a} = (1/\sqrt{2},\, 1/\sqrt{2},\, 0)$, $\mathbf{b} = (1/\sqrt{2},\, -1/\sqrt{2},\, 0)$, and $\mathbf{a} * \mathbf{b} = (1/2,\, -1/2,\, 0)$. Thus, $\mathbf{a} * \mathbf{b}$ does not transform like a vector, because its magnitude depends on the choice of axes. So, above definition is a bad one.

Consider, now, the *cross product* or *vector product*:

$$\mathbf{a} \times \mathbf{b} \equiv (a_y\, b_z - a_z\, b_y,\; a_z\, b_x - a_x\, b_z,\; a_x\, b_y - a_y\, b_x) = \mathbf{c}. \tag{A.34}$$

Does this rather unlikely combination transform like a vector? Let us try rotating the coordinate axes through an angle θ about Oz using Equations (A.20)–(A.22). In the new coordinate system,

$$\begin{aligned} c_{x'} &= (-a_x\, \sin\theta + a_y\, \cos\theta)\, b_z - a_z\, (-b_x\, \sin\theta + b_y\, \cos\theta) \\ &= (a_y\, b_z - a_z\, b_y)\, \cos\theta + (a_z\, b_x - a_x\, b_z)\, \sin\theta \\ &= c_x\, \cos\theta + c_y\, \sin\theta. \end{aligned} \tag{A.35}$$

Thus, the x-component of $\mathbf{a} \times \mathbf{b}$ transforms correctly. It can easily be shown that the other components transform correctly as well, and that all components also transform correctly under rotation about Ox and Oy. Thus, $\mathbf{a} \times \mathbf{b}$ is a proper vector. Incidentally, $\mathbf{a} \times \mathbf{b}$ is the *only* simple combination of the components of two vectors that transforms like a vector (which is non-coplanar with **a** and **b**). The cross product is *anticommutative*,

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}, \tag{A.36}$$

distributive,

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}, \tag{A.37}$$

but is *not* associative,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}. \tag{A.38}$$

The cross product transforms like a vector, which means that it must have a well-defined direction and magnitude. We can show that $\mathbf{a} \times \mathbf{b}$ is *perpendicular* to both **a** and **b**. Consider $\mathbf{a} \cdot \mathbf{a} \times \mathbf{b}$. If this is zero then the cross product must be perpendicular to **a**. Now,

$$\begin{aligned} \mathbf{a} \cdot \mathbf{a} \times \mathbf{b} &= a_x\, (a_y\, b_z - a_z\, b_y) + a_y\, (a_z\, b_x - a_x\, b_z) + a_z\, (a_x\, b_y - a_y\, b_x) \\ &= 0. \end{aligned} \tag{A.39}$$

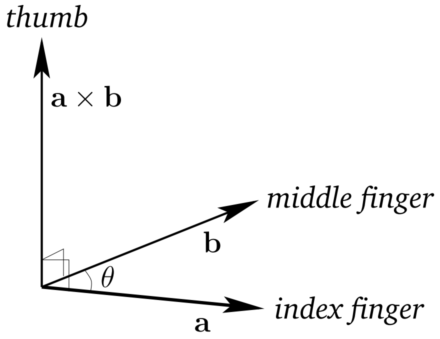


Figure A.7: *The right-hand rule for cross products. Here,* θ *is less that* 180°.

Therefore, $\mathbf{a}\times\mathbf{b}$ is perpendicular to $\mathbf{a}$. Likewise, it can be demonstrated that $\mathbf{a}\times\mathbf{b}$ is perpendicular to $\mathbf{b}$. The vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a}\times\mathbf{b}$ form a *right-handed* set, like the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$. In fact, $\mathbf{e}_x\times\mathbf{e}_y = \mathbf{e}_z$. This defines a unique direction for $\mathbf{a}\times\mathbf{b}$, which is obtained from a right-hand rule—see Figure A.7.

Let us now evaluate the magnitude of $\mathbf{a}\times\mathbf{b}$. We have

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})^2 &= (a_y\,b_z - a_z\,b_y)^2 + (a_z\,b_x - a_x\,b_z)^2 + (a_x\,b_y - a_y\,b_x)^2\\ &= (a_x^{\,2}+a_y^{\,2}+a_z^{\,2})\,(b_x^{\,2}+b_y^{\,2}+b_z^{\,2}) - (a_x\,b_x + a_y\,b_y + a_z\,b_z)^2\\ &= |\mathbf{a}|^2\,|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2\\ &= |\mathbf{a}|^2\,|\mathbf{b}|^2 - |\mathbf{a}|^2\,|\mathbf{b}|^2\,\cos^2\theta = |\mathbf{a}|^2\,|\mathbf{b}|^2\,\sin^2\theta. \end{aligned} \tag{A.40}$$

Thus,

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\,\sin\theta, \tag{A.41}$$

where θ is the angle subtended between $\mathbf{a}$ and $\mathbf{b}$. Clearly, $\mathbf{a}\times\mathbf{a} = \mathbf{0}$ for any vector, since θ is always zero in this case. Also, if $\mathbf{a}\times\mathbf{b} = \mathbf{0}$ then either $|\mathbf{a}| = 0$, $|\mathbf{b}| = 0$, or $\mathbf{b}$ is parallel (or antiparallel) to $\mathbf{a}$.

Consider the parallelogram defined by the vectors $\mathbf{a}$ and $\mathbf{b}$—see Figure A.8. The *scalar area* of the parallelogram is $a\,b\sin\theta$. By convention, the *vector area* has the magnitude of the scalar area, and is normal to the plane of the parallelogram, in the sense obtained from a right-hand circulation rule by rotating $\mathbf{a}$ on to $\mathbf{b}$ (through an acute angle): *i.e.*, if the fingers of the right-hand circulate in the direction of rotation then the thumb of the right-hand indicates the direction of the vector area. So, the vector area is coming out of the page in Figure A.8. It follows that

$$\mathbf{S} = \mathbf{a}\times\mathbf{b}, \tag{A.42}$$

Suppose that a force $\mathbf{F}$ is applied at position $\mathbf{r}$—see Figure A.9. The torque about the origin O is the product of the magnitude of the force and the length of the lever arm OQ. Thus, the magnitude of the torque is $|\mathbf{F}|\,|\mathbf{r}|\,\sin\theta$. The direction of the torque is conventionally defined as the direction of the axis through O about which the force tries to

rotate objects, in the sense determined by a right-hand circulation rule. Hence, the torque is out of the page in Figure A.9. It follows that the vector torque is given by

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}. \tag{A.43}$$

The angular momentum, $\mathbf{l}$, of a particle of linear momentum $\mathbf{p}$ and position vector $\mathbf{r}$ is simply defined as the moment of its momentum about the origin: *i.e.*,

$$\mathbf{l} = \mathbf{r} \times \mathbf{p}. \tag{A.44}$$

A.8 Rotation

Let us try to define a rotation vector $\boldsymbol{\theta}$ whose magnitude is the angle of the rotation, θ, and whose direction is parallel to the axis of rotation, in the sense determined by a right-hand circulation rule. Unfortunately, this is not a good vector. The problem is that the addition of rotations is not commutative, whereas vector addition is commuative. Figure A.10 shows the effect of applying two successive 90° rotations, one about Ox, and the other about the Oz, to a standard six-sided die. In the left-hand case, the z-rotation is applied before the x-rotation, and *vice versa* in the right-hand case. It can be seen that the die ends up in two completely different states. In other words, the z-rotation plus the x-rotation does not equal the x-rotation plus the z-rotation. This non-commuting algebra cannot be represented by vectors. So, although rotations have a well-defined magnitude and direction, they are *not* vector quantities.

But, this is not quite the end of the story. Suppose that we take a general vector $\mathbf{a}$ and rotate it about Oz by a *small* angle $\delta\theta_z$. This is equivalent to rotating the coordinate axes about the Oz by $-\delta\theta_z$. According to Equations (A.20)–(A.22), we have

$$\mathbf{a}' \simeq \mathbf{a} + \delta\theta_z\,\mathbf{e}_z \times \mathbf{a}, \tag{A.45}$$

where use has been made of the small angle approximations $\sin\theta \simeq \theta$ and $\cos\theta \simeq 1$. The above equation can easily be generalized to allow small rotations about Ox and Oy by $\delta\theta_x$ and $\delta\theta_y$, respectively. We find that

$$\mathbf{a}' \simeq \mathbf{a} + \delta\boldsymbol{\theta} \times \mathbf{a}, \tag{A.46}$$

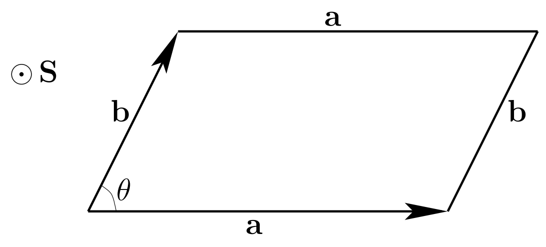


Figure A.8: *A vector parallelogram.*

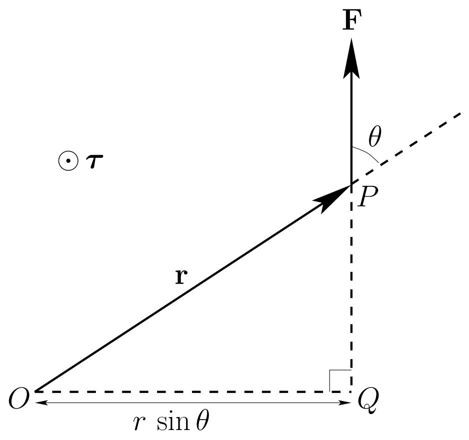


Figure A.9: *A torque.*

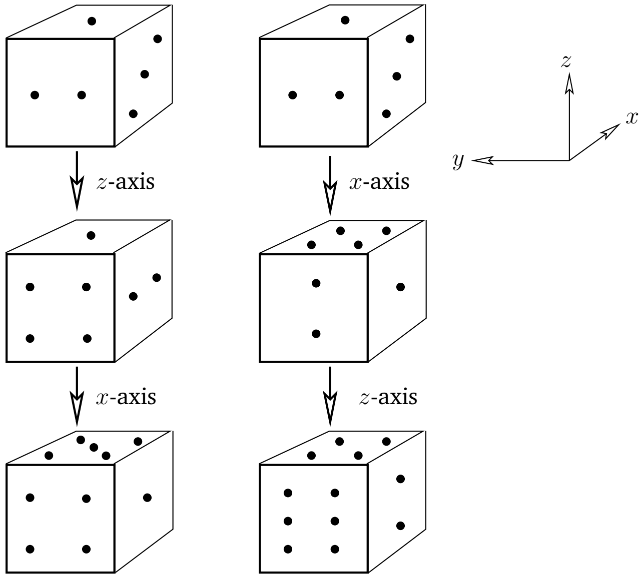


Figure A.10: *Effect of successive rotations about perpendicular axes on a six-sided die.*

where

$$\delta\boldsymbol{\theta} = \delta\theta_x\,\mathbf{e}_x + \delta\theta_y\,\mathbf{e}_y + \delta\theta_z\,\mathbf{e}_z. \tag{A.47}$$

Clearly, we can define a rotation vector, $\delta\boldsymbol{\theta}$, but it only works for *small* angle rotations (*i.e.*, sufficiently small that the small angle approximations of sine and cosine are good). According to the above equation, a small z-rotation plus a small x-rotation is (approximately) equal to the two rotations applied in the opposite order. The fact that infinitesimal rotation is a vector implies that angular velocity,

$$\boldsymbol{\omega} = \lim_{\delta t\to 0}\frac{\delta\boldsymbol{\theta}}{\delta t}, \tag{A.48}$$

must be a vector as well. Also, if $\mathbf{a}'$ is interpreted as $\mathbf{a}(t+\delta t)$ in Equation (A.46) then it follows that the equation of motion of a vector which precesses about the origin with some angular velocity $\boldsymbol{\omega}$ is

$$\frac{d\mathbf{a}}{dt} = \boldsymbol{\omega}\times\mathbf{a}. \tag{A.49}$$

A.9 Scalar Triple Product

Consider three vectors **a**, **b**, and **c**. The *scalar triple product* is defined $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$. Now, $\mathbf{b}\times\mathbf{c}$ is the vector area of the parallelogram defined by **b** and **c**. So, $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$ is the scalar area of this parallelogram multiplied by the component of **a** in the direction of its normal. It follows that $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}$ is the *volume* of the parallelepiped defined by vectors **a**, **b**, and **c**—see Figure A.11. This volume is independent of how the triple product is formed from **a**, **b**, and **c**, except that

$$\mathbf{a}\cdot\mathbf{b}\times\mathbf{c} = -\mathbf{a}\cdot\mathbf{c}\times\mathbf{b}. \tag{A.50}$$

So, the "volume" is positive if **a**, **b**, and **c** form a right-handed set (*i.e.*, if **a** lies above the plane of **b** and **c**, in the sense determined from a right-hand circulation rule by rotating **b** onto **c**) and negative if they form a left-handed set. The triple product is unchanged if the dot and cross product operators are interchanged,

$$\mathbf{a}\cdot\mathbf{b}\times\mathbf{c} = \mathbf{a}\times\mathbf{b}\cdot\mathbf{c}. \tag{A.51}$$

The triple product is also invariant under any cyclic permutation of **a**, **b**, and **c**,

$$\mathbf{a}\cdot\mathbf{b}\times\mathbf{c} = \mathbf{b}\cdot\mathbf{c}\times\mathbf{a} = \mathbf{c}\cdot\mathbf{a}\times\mathbf{b}, \tag{A.52}$$

but any anti-cyclic permutation causes it to change sign,

$$\mathbf{a}\cdot\mathbf{b}\times\mathbf{c} = -\mathbf{b}\cdot\mathbf{a}\times\mathbf{c}. \tag{A.53}$$

The scalar triple product is zero if any two of **a**, **b**, and **c** are parallel, or if **a**, **b**, and **c** are coplanar.

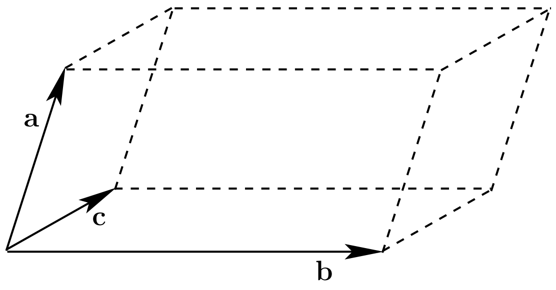


Figure A.11: *A vector parallelepiped.*

If **a**, **b**, and **c** are non-coplanar then any vector **r** can be written in terms of them: *i.e.*,

$$\mathbf{r} = \alpha\,\mathbf{a} + \beta\,\mathbf{b} + \gamma\,\mathbf{c}. \tag{A.54}$$

Forming the dot product of this equation with $\mathbf{b}\times\mathbf{c}$, we then obtain

$$\mathbf{r}\cdot\mathbf{b}\times\mathbf{c} = \alpha\,\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}, \tag{A.55}$$

so

$$\alpha = \frac{\mathbf{r}\cdot\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}. \tag{A.56}$$

Analogous expressions can be written for β and γ. The parameters α, β, and γ are uniquely determined provided $\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}\neq 0$: *i.e.*, provided that the three vectors are non-coplanar.

A.10 Vector Triple Product

For three vectors **a**, **b**, and **c**, the *vector triple product* is defined $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. The brackets are important because $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})\neq(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$. In fact, it can be demonstrated that

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) \equiv (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c} \tag{A.57}$$

and

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c} \equiv (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{b}\cdot\mathbf{c})\,\mathbf{a}. \tag{A.58}$$

Let us try to prove the first of the above theorems. The left-hand side and the right-hand side are both proper vectors, so if we can prove this result in one particular coordinate system then it must be true in general. Let us take convenient axes such that Ox lies along **b**, and **c** lies in the x-y plane. It follows that $\mathbf{b}\equiv(b_x,\,0,\,0)$, $\mathbf{c}\equiv(c_x,\,c_y,\,0)$, and $\mathbf{a}\equiv(a_x,\,a_y,\,a_z)$. The vector $\mathbf{b}\times\mathbf{c}$ is directed along Oz: *i.e.*, $\mathbf{b}\times\mathbf{c}\equiv(0,\,0,\,b_x\,c_y)$. Hence, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ lies in the x-y plane: *i.e.*, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})\equiv(a_y\,b_x\,c_y,\,-a_x\,b_x\,c_y,\,0)$. This is the left-hand side of Equation (A.57) in our convenient coordinate system. To evaluate the right-hand side, we need $\mathbf{a}\cdot\mathbf{c}=a_x\,c_x+a_y\,c_y$ and $\mathbf{a}\cdot\mathbf{b}=a_x\,b_x$. It follows that the right-hand side is

$$\begin{aligned}\text{RHS} &= (\,[a_x\,c_x+a_y\,c_y]\,b_x,\,0,\,0) - (a_x\,b_x\,c_x,\,a_x\,b_x\,c_y,\,0)\\ &= (a_y\,c_y\,b_x,\,-a_x\,b_x\,c_y,\,0) = \text{LHS},\end{aligned} \tag{A.59}$$

which proves the theorem.

A.11 Vector Calculus

Suppose that vector **a** varies with time, so that $\mathbf{a} = \mathbf{a}(t)$. The time derivative of the vector is defined

$$\frac{d\mathbf{a}}{dt} = \lim_{\delta t \to 0} \left[\frac{\mathbf{a}(t + \delta t) - \mathbf{a}(t)}{\delta t} \right]. \tag{A.60}$$

When written out in component form this becomes

$$\frac{d\mathbf{a}}{dt} \equiv \left(\frac{da_x}{dt}, \frac{da_y}{dt}, \frac{da_z}{dt} \right). \tag{A.61}$$

Suppose that **a** is, in fact, the product of a scalar $\phi(t)$ and another vector $\mathbf{b}(t)$. What now is the time derivative of **a**? We have

$$\frac{da_x}{dt} = \frac{d}{dt}(\phi\, b_x) = \frac{d\phi}{dt}\, b_x + \phi\, \frac{db_x}{dt}, \tag{A.62}$$

which implies that

$$\frac{d\mathbf{a}}{dt} = \frac{d\phi}{dt}\,\mathbf{b} + \phi\,\frac{d\mathbf{b}}{dt}. \tag{A.63}$$

Moreover, it is easily demonstrated that

$$\frac{d}{dt}\,(\mathbf{a}\cdot\mathbf{b}) = \frac{d\mathbf{a}}{dt}\cdot\mathbf{b} + \mathbf{a}\cdot\frac{d\mathbf{b}}{dt}, \tag{A.64}$$

and

$$\frac{d}{dt}\,(\mathbf{a}\times\mathbf{b}) = \frac{d\mathbf{a}}{dt}\times\mathbf{b} + \mathbf{a}\times\frac{d\mathbf{b}}{dt}. \tag{A.65}$$

Hence, it can be seen that the laws of vector differentiation are analogous to those in conventional calculus.

A.12 Line Integrals

Consider a two-dimensional function $f(x, y)$ which is defined for all x and y. What is meant by the integral of f along a given curve joining the points P and Q in the x-y plane? Well, we first draw out f as a function of length l along the path—see Figure A.12. The integral is then simply given by

$$\int_P^Q f(x, y)\, dl = \text{Area under the curve}, \tag{A.66}$$

where $dl = \sqrt{dx^2 + dy^2}$.

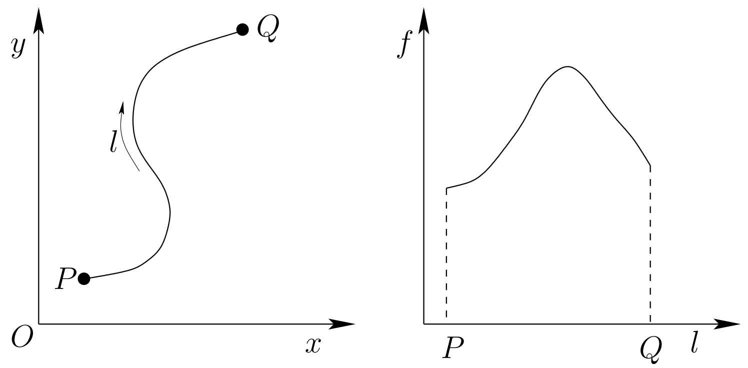


Figure A.12: *A line integral.*

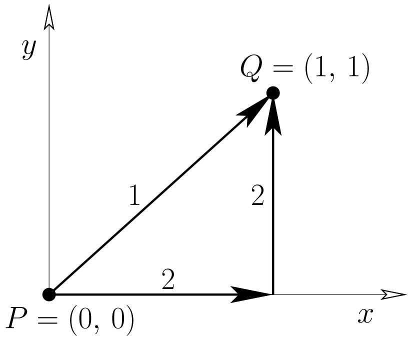


Figure A.13: *An example line integral.*

As an example of this, consider the integral of $f(x, y) = x\,y^2$ between P and Q along the two routes indicated in Figure A.13. Along route 1 we have $x = y$, so $dl = \sqrt{2}\,dx$. Thus,

$$\int_P^Q x\,y^2\,dl = \int_0^1 x^3\,\sqrt{2}\,dx = \frac{\sqrt{2}}{4}. \tag{A.67}$$

The integration along route 2 gives

$$\begin{aligned}\int_P^Q x\,y^2\,dl &= \int_0^1 x\,y^2\,dx\bigg|_{y=0} + \int_0^1 x\,y^2\,dy\bigg|_{x=1} \\ &= 0 + \int_0^1 y^2\,dy = \frac{1}{3}.\end{aligned} \tag{A.68}$$

Note that the integral depends on the route taken between the initial and final points.

The most common type of line integral is that in which the contributions from dx and dy are evaluated separately, rather that through the path length dl: *i.e.*,

$$\int_P^Q [f(x, y)\,dx + g(x, y)\,dy]\,. \tag{A.69}$$

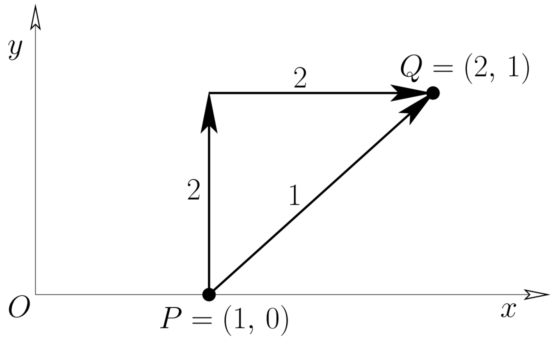


Figure A.14: *An example line integral.*

As an example of this, consider the integral

$$\int_P^Q \left[y\,dx + x^3\,dy\right] \tag{A.70}$$

along the two routes indicated in Figure A.14. Along route 1 we have $x = y + 1$ and $dx = dy$, so

$$\int_P^Q \left[y\,dx + x^3\,dy\right] = \int_0^1 \left[y\,dy + (y+1)^3\,dy\right] = \frac{17}{4}. \tag{A.71}$$

Along route 2,

$$\int_P^Q \left[y\,dx + x^3\,dy\right] = \int_0^1 x^3\,dy\bigg|_{x=1} + \int_1^2 y\,dx\bigg|_{y=1} = \frac{7}{4}. \tag{A.72}$$

Again, the integral depends on the path of integration.

Suppose that we have a line integral which does *not* depend on the path of integration. It follows that

$$\int_P^Q (f\,dx + g\,dy) = F(Q) - F(P) \tag{A.73}$$

for some function F. Given $F(P)$ for one point P in the x-y plane, then

$$F(Q) = F(P) + \int_P^Q (f\,dx + g\,dy) \tag{A.74}$$

defines $F(Q)$ for all other points in the plane. We can then draw a contour map of $F(x,y)$. The line integral between points P and Q is simply the change in height in the contour map between these two points:

$$\int_P^Q (f\,dx + g\,dy) = \int_P^Q dF(x,y) = F(Q) - F(P). \tag{A.75}$$

Thus,

$$dF(x,y) = f(x,y)\,dx + g(x,y)\,dy. \tag{A.76}$$

For instance, if $F = x^3\,y$ then $dF = 3\,x^2\,y\,dx + x^3\,dy$ and

$$\int_P^Q \left(3\,x^2\,y\,dx + x^3\,dy\right) = \left[x^3\,y\right]_P^Q \tag{A.77}$$

is independent of the path of integration.

It is clear that there are two distinct types of line integral. Those which depend only on their endpoints and not on the path of integration, and those which depend both on their endpoints and the integration path. Later on, we shall learn how to distinguish between these two types (see Section A.15).

A.13 Vector Line Integrals

A *vector field* is defined as a set of vectors associated with each point in space. For instance, the velocity $\mathbf{v}(\mathbf{r})$ in a moving liquid (*e.g.*, a whirlpool) constitutes a vector field. By analogy, a *scalar field* is a set of scalars associated with each point in space. An example of a scalar field is the temperature distribution $T(\mathbf{r})$ in a furnace.

Consider a general vector field $\mathbf{A}(\mathbf{r})$. Let $d\mathbf{r} \equiv (dx,\, dy,\, dz)$ be the vector element of line length. Vector line integrals often arise as

$$\int_P^Q \mathbf{A}\cdot d\mathbf{r} = \int_P^Q (A_x\,dx + A_y\,dy + A_z\,dz). \tag{A.78}$$

For instance, if $\mathbf{A}$ is a force-field then the line integral is the work done in going from P to Q.

As an example, consider the work done by a repulsive inverse-square central field, $\mathbf{F} = -\mathbf{r}/|r^3|$. The element of work done is $dW = \mathbf{F}\cdot d\mathbf{r}$. Take $P = (\infty, 0, 0)$ and $Q = (a, 0, 0)$. Route 1 is along the x-axis, so

$$W = \int_\infty^a \left(-\frac{1}{x^2}\right) dx = \left[\frac{1}{x}\right]_\infty^a = \frac{1}{a}. \tag{A.79}$$

The second route is, firstly, around a large circle ($r =$ constant) to the point (a, ∞, 0), and then parallel to the y-axis—see Figure A.15. In the first part, no work is done, since $\mathbf{F}$ is perpendicular to $d\mathbf{r}$. In the second part,

$$W = \int_\infty^0 \frac{-y\,dy}{(a^2+y^2)^{3/2}} = \left[\frac{1}{(y^2+a^2)^{1/2}}\right]_\infty^0 = \frac{1}{a}. \tag{A.80}$$

In this case, the integral is independent of the path. However, not all vector line integrals are path independent.

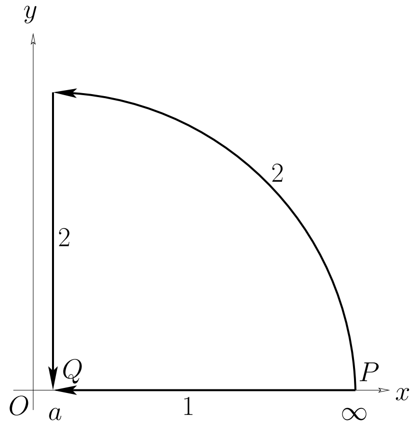


Figure A.15: *An example vector line integral.*

A.14 Volume Integrals

A volume integral takes the form

$$\iiint_V f(x, y, z)\, dV, \tag{A.81}$$

where V is some volume, and $dV = dx\, dy\, dz$ is a small volume element. The volume element is sometimes written $d^3\mathbf{r}$, or even $d\tau$.

As an example of a volume integral, let us evaluate the center of gravity of a solid pyramid. Suppose that the pyramid has a square base of side a, a height a, and is composed of material of uniform density. Let the centroid of the base lie at the origin, and let the apex lie at $(0, 0, a)$. By symmetry, the center of mass lies on the line joining the centroid to the apex. In fact, the height of the center of mass is given by

$$\overline{z} = \iiint z\, dV \bigg/ \iiint dV. \tag{A.82}$$

The bottom integral is just the volume of the pyramid, and can be written

$$\begin{aligned}\iiint dV &= \int_0^a dz \int_{-(a-z)/2}^{(a-z)/2} dy \int_{-(a-z)/2}^{(a-z)/2} dx = \int_0^a (a-z)^2\, dz = \int_0^a (a^2 - 2\, a\, z + z^2)\, dz \\ &= \left[a^2\, z - a\, z^2 + z^3/3\right]_0^a = \frac{1}{3}\, a^3. \end{aligned} \tag{A.83}$$

Here, we have evaluated the z-integral last because the limits of the x- and y- integrals are z-dependent. The top integral takes the form

$$\begin{aligned}\iiint z\, dV &= \int_0^a z\, dz \int_{-(a-z)/2}^{(a-z)/2} dy \int_{-(a-z)/2}^{(a-z)/2} dx = \int_0^a z\, (a-z)^2\, dz = \int_0^a (z\, a^2 - 2\, a\, z^2 + z^3)\, dz \\ &= \left[a^2\, z^2/2 - 2\, a\, z^3/3 + z^4/4\right]_0^a = \frac{1}{12}\, a^4. \end{aligned} \tag{A.84}$$

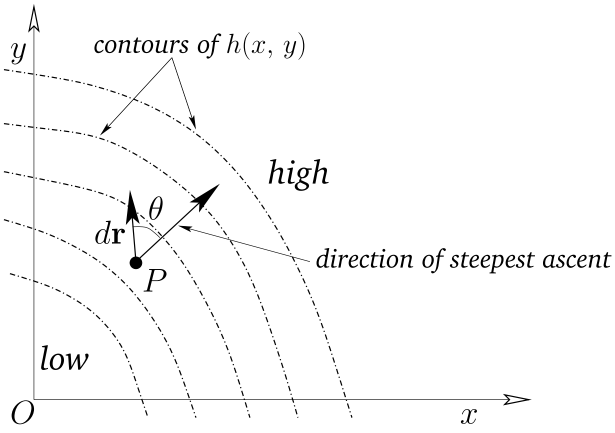


Figure A.16: *A two-dimensional gradient.*

Thus,

$$\bar{z} = \frac{1}{12}\,a^4 \Big/ \frac{1}{3}\,a^3 = \frac{1}{4}\,a. \tag{A.85}$$

In other words, the center of mass of a pyramid lies one quarter of the way between the centroid of the base and the apex.

A.15 Gradient

A one-dimensional function $f(x)$ has a gradient df/dx which is defined as the slope of the tangent to the curve at x. We wish to extend this idea to cover scalar fields in two and three dimensions.

Consider a two-dimensional scalar field $h(x, y)$, which is (say) height above sea-level in a hilly region. Let $d\mathbf{r} \equiv (dx, dy)$ be an element of horizontal distance. Consider dh/dr, where dh is the change in height after moving an infinitesimal distance $d\mathbf{r}$. This quantity is somewhat like the one-dimensional gradient, except that dh depends on the *direction* of $d\mathbf{r}$, as well as its magnitude. In the immediate vicinity of some point P, the slope reduces to an inclined plane—see Figure A.16. The largest value of dh/dr is straight up the slope. It is easily shown that for any other direction

$$\frac{dh}{dr} = \left(\frac{dh}{dr}\right)_{\rm max} \cos\theta, \tag{A.86}$$

where θ is the angle shown in Figure A.16. Let us define a two-dimensional vector, $\mathbf{grad}\,h$, called the *gradient* of h, whose magnitude is $(dh/dr)_{\rm max}$, and whose direction is the direction of steepest ascent. The $\cos\theta$ variation exhibited in the above expression ensures that the component of $\mathbf{grad}\,h$ in any direction is equal to dh/dr for that direction.

The component of dh/dr in the x-direction can be obtained by plotting out the profile of h at constant y, and then finding the slope of the tangent to the curve at given x. This

quantity is known as the *partial derivative* of h with respect to x at constant y, and is denoted $(\partial h/\partial x)_y$. Likewise, the gradient of the profile at constant x is written $(\partial h/\partial y)_x$. Note that the subscripts denoting constant-x and constant-y are usually omitted, unless there is any ambiguity. If follows that in component form

$$\mathbf{grad}\, h \equiv \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}\right). \tag{A.87}$$

Now, the equation of the tangent plane at $P = (x_0, y_0)$ is

$$h_T(x, y) = h(x_0, y_0) + \alpha\,(x - x_0) + \beta\,(y - y_0). \tag{A.88}$$

This has the same local gradients as $h(x, y)$, so

$$\alpha = \frac{\partial h}{\partial x}, \qquad \beta = \frac{\partial h}{\partial y}, \tag{A.89}$$

by differentiation of the above. For small $dx = x - x_0$ and $dy = y - y_0$, the function h is coincident with the tangent plane, so

$$dh = \frac{\partial h}{\partial x}\, dx + \frac{\partial h}{\partial y}\, dy. \tag{A.90}$$

But, $\mathbf{grad}\, h \equiv (\partial h/\partial x, \partial h/\partial y)$ and $d\mathbf{r} \equiv (dx, dy)$, so

$$dh = \mathbf{grad}\, h \cdot d\mathbf{r}. \tag{A.91}$$

Incidentally, the above equation demonstrates that $\mathbf{grad}\, h$ is a proper vector, since the left-hand side is a scalar, and, according to the properties of the dot product, the right-hand side is also a scalar provided that $d\mathbf{r}$ and $\mathbf{grad}\, h$ are both proper vectors ($d\mathbf{r}$ is an obvious vector, because it is directly derived from displacements).

Consider, now, a three-dimensional temperature distribution $T(x, y, z)$ in (say) a reaction vessel. Let us define $\mathbf{grad}\, T$, as before, as a vector whose magnitude is $(dT/dr)_{\rm max}$, and whose direction is the direction of the maximum gradient. This vector is written in component form

$$\mathbf{grad}\, T \equiv \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right). \tag{A.92}$$

Here, $\partial T/\partial x \equiv (\partial T/\partial x)_{y,z}$ is the gradient of the one-dimensional temperature profile at constant y and z. The change in T in going from point P to a neighbouring point offset by $d\mathbf{r} \equiv (dx, dy, dz)$ is

$$dT = \frac{\partial T}{\partial x}\, dx + \frac{\partial T}{\partial y}\, dy + \frac{\partial T}{\partial z}\, dz. \tag{A.93}$$

In vector form, this becomes

$$dT = \mathbf{grad}\, T \cdot d\mathbf{r}. \tag{A.94}$$

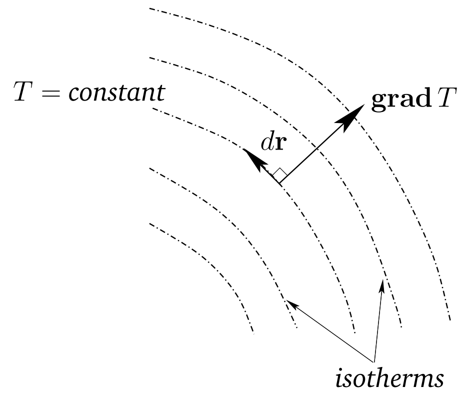


Figure A.17: *Isotherms.*

Suppose that $dT = 0$ for some $d\mathbf{r}$. It follows that

$$dT = \mathbf{grad}\, T \cdot d\mathbf{r} = 0. \tag{A.95}$$

So, $d\mathbf{r}$ is perpendicular to $\mathbf{grad}\, T$. Since $dT = 0$ along so-called "isotherms" (*i.e.*, contours of the temperature), we conclude that the isotherms (contours) are everywhere perpendicular to $\mathbf{grad}\, T$—see Figure A.17.

It is, of course, possible to integrate dT. For instance, the line integral of dT between points P and Q is written

$$\int_P^Q dT = \int_P^Q \mathbf{grad}\, T \cdot d\mathbf{r} = T(Q) - T(P). \tag{A.96}$$

This integral is clearly independent of the path taken between P and Q, so $\int_P^Q \mathbf{grad}\, T \cdot d\mathbf{r}$ must be path independent.

Consider a vector field $\mathbf{A}(\mathbf{r})$. In general, the line integral $\int_P^Q \mathbf{A} \cdot d\mathbf{r}$ depends on the path taken between the end points, but for some special vector fields the integral is path independent. Such fields are called *conservative* fields. It can be shown that if $\mathbf{A}$ is a conservative field then $\mathbf{A} = \mathbf{grad}\, V$ for some scalar field V. The proof of this is straightforward. Keeping P fixed, we have

$$\int_P^Q \mathbf{A} \cdot d\mathbf{r} = V(Q), \tag{A.97}$$

where $V(Q)$ is a well-defined function, due to the path independent nature of the line integral. Consider moving the position of the end point by an infinitesimal amount dx in the x-direction. We have

$$V(Q + dx) = V(Q) + \int_Q^{Q+dx} \mathbf{A} \cdot d\mathbf{r} = V(Q) + A_x\, dx. \tag{A.98}$$

Hence,

$$\frac{\partial V}{\partial x} = A_x, \tag{A.99}$$

with analogous relations for the other components of $\mathbf{A}$. It follows that

$$\mathbf{A} = \mathbf{grad}\, V. \tag{A.100}$$

In Newtonian dynamics, the force due to gravity is a good example of a conservative field. Now, if $\mathbf{A}(\mathbf{r})$ is a force-field then $\int \mathbf{A}\cdot d\mathbf{r}$ is the work done in traversing some path. If $\mathbf{A}$ is conservative then

$$\oint \mathbf{A}\cdot d\mathbf{r} = 0, \tag{A.101}$$

where $\oint$ corresponds to the line integral around a *closed* loop. The fact that zero net work is done in going around a closed loop is equivalent to the conservation of energy (which is why conservative fields are called "conservative"). A good example of a non-conservative field is the force due to friction. Clearly, a frictional system loses energy in going around a closed cycle, so $\oint \mathbf{A}\cdot d\mathbf{r} \neq 0$.

A.16 Grad Operator

It is useful to define the vector *operator*

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right), \tag{A.102}$$

which is usually called the *grad* or *del* operator. This operator acts on everything to its right in a expression, until the end of the expression or a closing bracket is reached. For instance,

$$\mathbf{grad}\, f = \nabla f \equiv \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right). \tag{A.103}$$

For two scalar fields ϕ and ψ,

$$\mathbf{grad}\,(\phi\,\psi) = \phi\ \mathbf{grad}\,\psi + \psi\ \mathbf{grad}\,\phi \tag{A.104}$$

can be written more succinctly as

$$\nabla(\phi\,\psi) = \phi\,\nabla\psi + \psi\,\nabla\phi. \tag{A.105}$$

Suppose that we rotate the coordinate axes through an angle θ about Oz. By analogy with Equations (A.17)–(A.19), the old coordinates $(x,\ y,\ z)$ are related to the new ones $(x',\ y',\ z')$ via

$$x = x'\cos\theta - y'\sin\theta, \tag{A.106}$$

$$y = x'\sin\theta + y'\cos\theta, \tag{A.107}$$

$$z = z'. \tag{A.108}$$

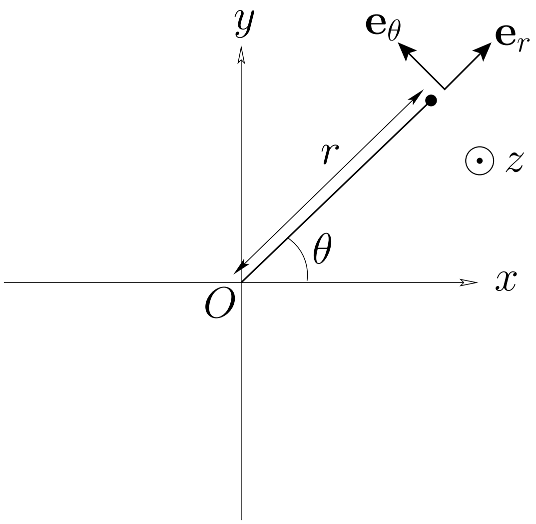


Figure A.18: *Cylindrical polar coordinates.*

Now,

$$\frac{\partial}{\partial x'} = \left(\frac{\partial x}{\partial x'}\right)_{y',z'} \frac{\partial}{\partial x} + \left(\frac{\partial y}{\partial x'}\right)_{y',z'} \frac{\partial}{\partial y} + \left(\frac{\partial z}{\partial x'}\right)_{y',z'} \frac{\partial}{\partial z}, \tag{A.109}$$

giving

$$\frac{\partial}{\partial x'} = \cos\theta \, \frac{\partial}{\partial x} + \sin\theta \, \frac{\partial}{\partial y}, \tag{A.110}$$

and

$$\nabla_{x'} = \cos\theta \, \nabla_x + \sin\theta \, \nabla_y. \tag{A.111}$$

It can be seen, from Equations (A.20)–(A.22), that the differential operator ∇ transforms in an analogous manner to a vector. This is another proof that ∇f is a good vector.

A.17 Curvilinear Coordinates

In the *cylindrical* coordinate system, the Cartesian coordinates x and y are replaced by $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$. Here, r is the perpendicular distance from the z-axis, and θ the angle subtended between the perpendicular radius vector and the x-axis—see Figure A.18. A general vector **A** is thus written

$$\mathbf{A} = A_r \, \mathbf{e}_r + A_\theta \, \mathbf{e}_\theta + A_z \, \mathbf{e}_z, \tag{A.112}$$

where $\mathbf{e}_r = \nabla r/|\nabla r|$ and $\mathbf{e}_\theta = \nabla\theta/|\nabla\theta|$—see Figure A.18. Note that the unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$, and $\mathbf{e}_z$ are mutually orthogonal. Hence, $A_r = \mathbf{A}\cdot\mathbf{e}_r$, *etc.* The volume element in this coordinate system is $d^3\mathbf{r} = r\,dr\,d\theta\,dz$. Moreover, the gradient of a general scalar field $V(\mathbf{r})$ takes the form

$$\nabla V = \frac{\partial V}{\partial r} \, \mathbf{e}_r + \frac{1}{r}\frac{\partial V}{\partial \theta} \, \mathbf{e}_\theta + \frac{\partial V}{\partial z} \, \mathbf{e}_z. \tag{A.113}$$

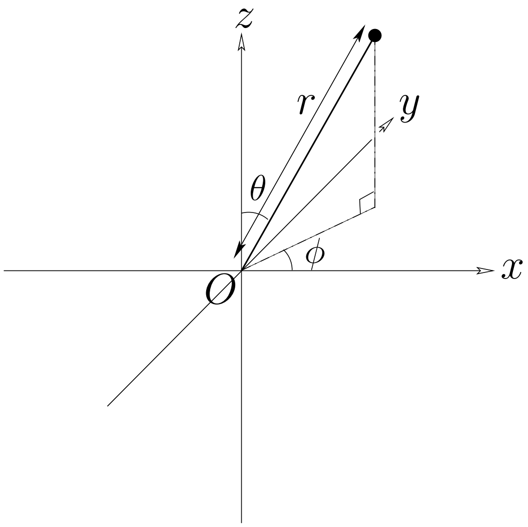


Figure A.19: *Spherical polar coordinates.*

In the *spherical* coordinate system, the Cartesian coordinates x, y, and z are replaced by $r = \sqrt{x^2 + y^2 + z^2}$, $\theta = \cos^{-1}(z/r)$, and $\phi = \tan^{-1}(y/x)$. Here, r is the radial distance from the origin, θ the angle subtended between the radius vector and the z-axis, and ϕ the angle subtended between the projection of the radius vector onto the x-y plane and the x-axis—see Figure A.19. Note that r and θ in the spherical system are *not* the same as their counterparts in the cylindrical system. A general vector $\mathbf{A}$ is written

$$\mathbf{A} = A_r\, \mathbf{e}_r + A_\theta\, \mathbf{e}_\theta + A_\phi\, \mathbf{e}_\phi, \tag{A.114}$$

where $\mathbf{e}_r = \nabla r/|\nabla r|$, $\mathbf{e}_\theta = \nabla\theta/|\nabla\theta|$, and $\mathbf{e}_\phi = \nabla\phi/|\nabla\phi|$. The unit vectors $\mathbf{e}_r$, $\mathbf{e}_\theta$, and $\mathbf{e}_\phi$ are mutually orthogonal. Hence, $A_r = \mathbf{A}\cdot\mathbf{e}_r$, *etc.* The volume element in this coordinate system is $d^3\mathbf{r} = r^2 \sin\theta\, dr\, d\theta\, d\phi$. Moreover, the gradient of a general scalar field $V(\mathbf{r})$ takes the form

$$\nabla V = \frac{\partial V}{\partial r}\,\mathbf{e}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\,\mathbf{e}_\theta + \frac{1}{r\,\sin\theta}\,\frac{\partial V}{\partial \phi}\,\mathbf{e}_\phi. \tag{A.115}$$

A.18 Exercises

A.1. The position vectors of the four points A, B, C, and D are $\mathbf{a}$, $\mathbf{b}$, $3\,\mathbf{a} + 2\,\mathbf{b}$, and $\mathbf{a} - 3\,\mathbf{b}$, respectively. Express $\overrightarrow{AC}$, $\overrightarrow{DB}$, $\overrightarrow{BC}$, and $\overrightarrow{CD}$ in terms of $\mathbf{a}$ and $\mathbf{b}$.

A.2. Prove the trigonometric law of sines

$$\frac{\sin a}{A} = \frac{\sin b}{B} = \frac{\sin c}{C}$$

using vector methods. Here, a, b, and c are the three angles of a plane triangle, and A, B, and C the lengths of the corresponding opposite sides.

A.3. Demonstrate using vectors that the diagonals of a parallelogram bisect one another. In addition, show that if the diagonals of a quadrilateral bisect one another then it is a parallelogram.

A.4. From the inequality

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos\theta \leq |\mathbf{a}|\,|\mathbf{b}|$$

deduce the triangle inequality

$$|\mathbf{a}+\mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|.$$

A.5. Find the scalar product $\mathbf{a}\cdot\mathbf{b}$ and the vector product $\mathbf{a}\times\mathbf{b}$ when

(a) $\mathbf{a} = \mathbf{e}_x + 3\,\mathbf{e}_y - \mathbf{e}_z, \quad \mathbf{b} = 3\,\mathbf{e}_x + 2\,\mathbf{e}_y + \mathbf{e}_z,$

(b) $\mathbf{a} = \mathbf{e}_x - 2\,\mathbf{e}_y + \mathbf{e}_z, \quad \mathbf{b} = 2\,\mathbf{e}_x + \mathbf{e}_y + \mathbf{e}_z.$

A.6. Which of the following statements regarding the three general vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are true?

(a) $\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b}) = (\mathbf{b}\times\mathbf{a})\cdot\mathbf{c}$.

(b) $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\times\mathbf{b})\times\mathbf{c}$.

(c) $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}$.

(d) $\mathbf{d} = \lambda\,\mathbf{a} + \mu\,\mathbf{b}$ implies that $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{d} = 0$.

(e) $\mathbf{a}\times\mathbf{c} = \mathbf{b}\times\mathbf{c}$ implies that $\mathbf{c}\cdot\mathbf{a} - \mathbf{c}\cdot\mathbf{b} = c\,|\mathbf{a}-\mathbf{b}|$.

(f) $(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{b}) = [\mathbf{b}\cdot(\mathbf{c}\times\mathbf{a})]\,\mathbf{b}$.

A.7. Prove that the length of the shortest straight-line from point $\mathbf{a}$ to the straight-line joining points $\mathbf{b}$ and $\mathbf{c}$ is

$$\frac{|\mathbf{a}\times\mathbf{b}+\mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}|}{|\mathbf{b}-\mathbf{c}|}.$$

A.8. Identify the following surfaces:

(a) $|\mathbf{r}| = a$,

(b) $\mathbf{r}\cdot\mathbf{n} = b$,

(c) $\mathbf{r}\cdot\mathbf{n} = c\,|\mathbf{r}|$,

(d) $|\mathbf{r} - (\mathbf{r}\cdot\mathbf{n})\,\mathbf{n}| = d$.

Here, $\mathbf{r}$ is the position vector, a, b, c, and d are positive constants, and $\mathbf{n}$ is a fixed unit vector.

A.9. Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be coplanar vectors related via

$$\alpha\,\mathbf{a} + \beta\,\mathbf{b} + \gamma\,\mathbf{c} = \mathbf{0},$$

where α, β, and γ are not all zero. Show that the condition for the points with position vectors $u\,\mathbf{a}$, $v\,\mathbf{b}$, and $w\,\mathbf{c}$ to be colinear is

$$\frac{\alpha}{u} + \frac{\beta}{v} + \frac{\gamma}{w} = 0.$$

A.10. If $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ are any vectors, demonstrate that $\mathbf{a} = \mathbf{q} + \lambda\,\mathbf{r}$, $\mathbf{b} = \mathbf{r} + \mu\,\mathbf{p}$, and $\mathbf{c} = \mathbf{p} + \nu\,\mathbf{q}$ are coplanar provided that $\lambda\,\mu\,\nu = -1$, where λ, μ, and ν are scalars. Show that this condition is satisfied when $\mathbf{a}$ is perpendicular to $\mathbf{p}$, $\mathbf{b}$ to $\mathbf{q}$, and $\mathbf{c}$ to $\mathbf{r}$.

A.11. The vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are non-coplanar, and form a non-orthogonal vector base. The vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, defined by

$$\mathbf{A} = \frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},$$

plus cyclic permutations, are said to be *reciprocal vectors*. Show that

$$\mathbf{a} = (\mathbf{B}\times\mathbf{C})/(\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}),$$

plus cyclic permutations.

A.12. In the notation of the previous question, demonstrate that the plane passing through points $\mathbf{a}/\alpha$, $\mathbf{b}/\beta$, and $\mathbf{c}/\gamma$ is normal to the direction of the vector

$$\mathbf{h} = \alpha\,\mathbf{A} + \beta\,\mathbf{B} + \gamma\,\mathbf{C}.$$

In addition, show that the perpendicular distance of the plane from the origin is $|\mathbf{h}|^{-1}$.

A.13. Evaluate $\oint \mathbf{A}\cdot d\mathbf{r}$ for

$$\mathbf{A} = \frac{x\,\mathbf{e}_x + y\,\mathbf{e}_y}{\sqrt{x^2+y^2}}$$

around the square whose sides are $x = 0$, $x = a$, $y = 0$, $y = a$.

A.14. Find the gradients of the following scalar functions of the position vector $\mathbf{r} = (x,\,y,\,z)$:

(a) $\mathbf{k}\cdot\mathbf{r}$,

(b) $|\mathbf{r}|^n$,

(c) $|\mathbf{r}-\mathbf{k}|^{-n}$,

(d) $\cos(\mathbf{k}\cdot\mathbf{r})$.

Here, $\mathbf{k}$ is a fixed vector.

B Useful Mathematrics

B.1 Conic Sections

The ellipse, the parabola, and the hyperbola are collectively known as *conic sections*, since these three types of curve can be obtained by taking various different plane sections of a right cone.

An *ellipse*, centered on the origin, of major radius a and minor radius b, which are aligned along the x- and y-axes, respectively (see Figure B.1), satisfies the following well-known equation:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{B.1}$$

Likewise, a parabola which is aligned along the $+x$-axis, and passes through the origin (see Figure B.2), satisfies:

$$y^2 - b\,x = 0, \tag{B.2}$$

where $b > 0$.

Finally, a hyperbola which is aligned along the $+x$-axis, and whose asymptotes intersect at the origin (see Figure B.3), satisfies:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{B.3}$$

Here, a is the distance of closest approach to the origin. The asymptotes subtend an angle $\phi = \tan^{-1}(b/a)$ with the x-axis.

It is not obvious, from the above formulae, what the ellipse, the parabola, and the hyperbola have in common. It turns out, in fact, that these three curves can all be represented as the locus of a movable point whose distance from a fixed point is in a constant ratio to its perpendicular distance to some fixed straight-line. Let the fixed point—which is termed the *focus* of the ellipse/parabola/hyperbola—lie at the origin, and let the fixed line—which is termed the *directrix*—correspond to $x = -d$ (with $d > 0$). Thus, the distance of a general point $(x,\ y)$ (which lies to the right of the directrix) from the focus is $r_1 = \sqrt{x^2 + y^2}$, whereas the perpendicular distance of the point from the directrix is $r_2 = x + d$—see Figure B.4. In polar coordinates, $r_1 = r$ and $r_2 = r\cos\theta + d$. Hence, the locus of a point for which r_1 and r_2 are in a fixed ratio satisfies the following equation:

$$\frac{r_1}{r_2} = \frac{\sqrt{x^2 + y^2}}{x + d} = \frac{r}{r\cos\theta + d} = e, \tag{B.4}$$

where $e \geq 0$ is a constant. When expressed in terms of polar coordinates, the above equation can be rearranged to give

$$r = \frac{r_c}{1 - e\cos\theta}, \tag{B.5}$$

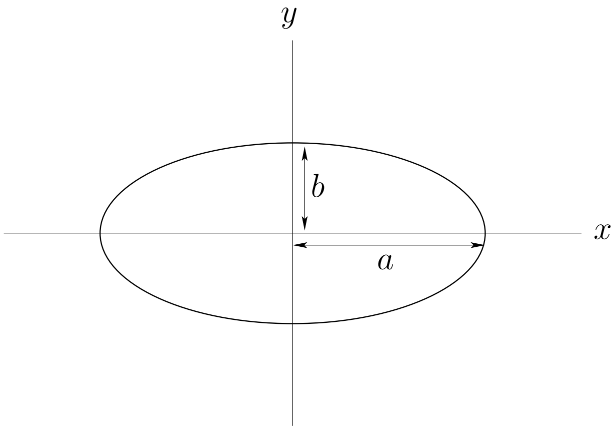


Figure B.1: *An ellipse.*

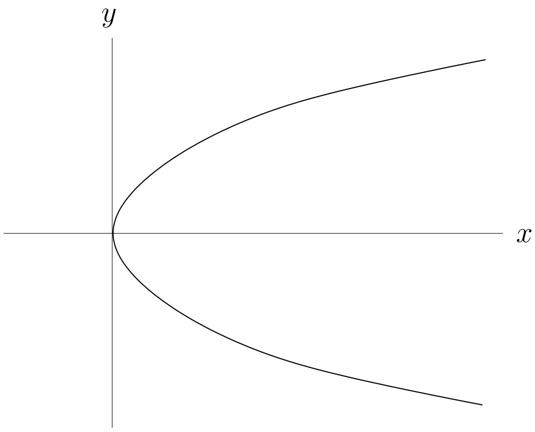


Figure B.2: *A parabola.*

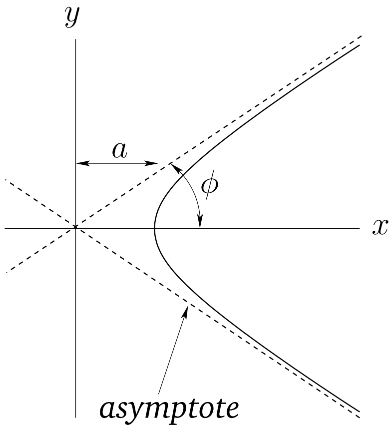


Figure B.3: *A hyperbola.*

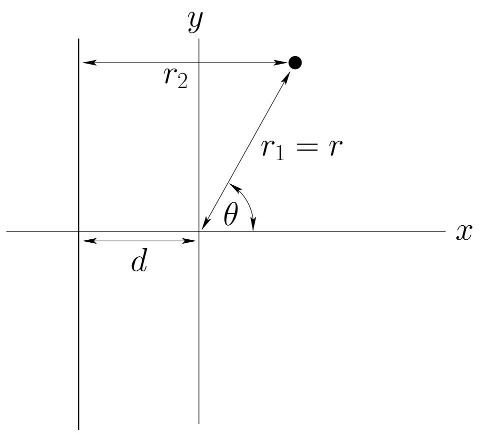


Figure B.4: *Conic sections in polar coordinates.*

where $r_c = e\,d$.

When written in terms of Cartesian coordinates, (B.4) can be rearranged to give

$$\frac{(x - x_c)^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{B.6}$$

for $e < 1$. Here,

$$a = \frac{r_c}{1 - e^2}, \tag{B.7}$$

$$b = \frac{r_c}{\sqrt{1 - e^2}} = \sqrt{1 - e^2}\,a, \tag{B.8}$$

$$x_c = \frac{e\,r_c}{1 - e^2} = e\,a. \tag{B.9}$$

Equation (B.6) can be recognized as the equation of an *ellipse* whose center lies at $(x_c,\ 0)$, and whose major and minor radii, a and b, are aligned along the x- and y-axes, respectively [*cf.*, Equation (B.1)]. Note, incidentally, that an ellipse actually possesses *two* focii located on the major axis ($y = 0$) a distance $e\,a$ on either side of the geometric center (*i.e.*, at $x = 0$ and $x = 2\,e\,a$).

When again written in terms of Cartesian coordinates, Equation (B.4) can be rearranged to give

$$y^2 - 2\,r_c\,(x - x_c) = 0, \tag{B.10}$$

for $e = 1$. Here, $x_c = -r_c/2$. This is the equation of a *parabola* which passes through the point $(x_c,\ 0)$, and which is aligned along the $+x$-direction [*cf.*, Equation (B.2)].

Finally, when written in terms of Cartesian coordinates, Equation (B.4) can be rearranged to give

$$\frac{(x - x_c)^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{B.11}$$

for $e > 1$. Here,

$$a = \frac{r_c}{e^2-1}, \tag{B.12}$$

$$b = \frac{r_c}{\sqrt{e^2-1}} = \sqrt{e^2-1}\,a, \tag{B.13}$$

$$x_c = -\frac{e\,r_c}{e^2-1} = -e\,a. \tag{B.14}$$

Equation (B.11) can be recognized as the equation of a *hyperbola* whose asymptotes intersect at $(x_c,\ 0)$, and which is aligned along the $+x$-direction. The asymptotes subtend an angle

$$\phi = \tan^{-1}\left(\frac{b}{a}\right) = \tan^{-1}(\sqrt{e^2-1}) \tag{B.15}$$

with the x-axis [*cf.*, Equation (B.3)].

In conclusion, Equation (B.5) is the polar equation of a *general conic section* which is *confocal with the origin* (*i.e.*, the origin lies at the focus). For $e < 1$, the conic section is an *ellipse*. For $e = 1$, the conic section is a *parabola*. Finally, for $e > 1$, the conic section is a *hyperbola*.

B.2 Matrix Eigenvalue Theory

Suppose that $\mathbf{A}$ is a *real symmetric* matrix of dimension n. If follows that $\mathbf{A}^* = \mathbf{A}$ and $\mathbf{A}^T = \mathbf{A}$, where * denotes a complex conjugate, and ^{T} denotes a transpose. Consider the matrix equation

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}. \tag{B.16}$$

Any column vector $\mathbf{x}$ which satisfies the above equation is called an *eigenvector* of $\mathbf{A}$. Likewise, the associated number λ is called an *eigenvalue* of $\mathbf{A}$. Let us investigate the properties of the eigenvectors and eigenvalues of a real symmetric matrix.

Equation (B.16) can be rearranged to give

$$(\mathbf{A} - \lambda\,\mathbf{1})\,\mathbf{x} = \mathbf{0}, \tag{B.17}$$

where $\mathbf{1}$ is the unit matrix. The above matrix equation is essentially a set of n homogeneous simultaneous algebraic equations for the n components of $\mathbf{x}$. A well-known property of such a set of equations is that it only has a non-trivial solution when the determinant of the associated matrix is set to zero. Hence, a necessary condition for the above set of equations to have a non-trivial solution is that

$$|\mathbf{A} - \lambda\,\mathbf{1}| = 0. \tag{B.18}$$

The above formula is essentially an nth-order *polynomial* equation for λ. We know that such an equation has n (possibly complex) roots. Hence, we conclude that there are n eigenvalues, and n associated eigenvectors, of the n-dimensional matrix $\mathbf{A}$.

Let us now demonstrate that the n eigenvalues and eigenvectors of the real symmetric matrix $\mathbf{A}$ are all *real*. We have

$$\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i, \tag{B.19}$$

and, taking the transpose and complex conjugate,

$$\mathbf{x}_i^{*\,T}\,\mathbf{A} = \lambda_i^*\,\mathbf{x}_i^{*\,T}, \tag{B.20}$$

where $\mathbf{x}_i$ and λ_i are the ith eigenvector and eigenvalue of $\mathbf{A}$, respectively. Left multiplying Equation (B.19) by $\mathbf{x}_i^{*\,T}$, we obtain

$$\mathbf{x}_i^{*\,T}\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i^{*\,T}\mathbf{x}_i. \tag{B.21}$$

Likewise, right multiplying (B.20) by $\mathbf{x}_i$, we get

$$\mathbf{x}_i^{*\,T}\,\mathbf{A}\,\mathbf{x}_i = \lambda_i^*\,\mathbf{x}_i^{*\,T}\mathbf{x}_i. \tag{B.22}$$

The difference of the previous two equations yields

$$(\lambda_i - \lambda_i^*)\,\mathbf{x}_i^{*\,T}\mathbf{x}_i = 0. \tag{B.23}$$

It follows that $\lambda_i = \lambda_i^*$, since $\mathbf{x}_i^{*\,T}\mathbf{x}_i$ (which is $\mathbf{x}_i^*\cdot\mathbf{x}_i$ in vector notation) is positive definite. Hence, λ_i is real. It immediately follows that $\mathbf{x}_i$ is real.

Next, let us show that two eigenvectors corresponding to two *different* eigenvalues are *mutually orthogonal*. Let

$$\mathbf{A}\,\mathbf{x}_i = \lambda_i\,\mathbf{x}_i, \tag{B.24}$$

$$\mathbf{A}\,\mathbf{x}_j = \lambda_j\,\mathbf{x}_j, \tag{B.25}$$

where $\lambda_i \neq \lambda_j$. Taking the transpose of the first equation and right multiplying by $\mathbf{x}_j$, and left multiplying the second equation by $\mathbf{x}_i^T$, we obtain

$$\mathbf{x}_i^T\,\mathbf{A}\,\mathbf{x}_j = \lambda_i\,\mathbf{x}_i^T\mathbf{x}_j, \tag{B.26}$$

$$\mathbf{x}_i^T\,\mathbf{A}\,\mathbf{x}_j = \lambda_j\,\mathbf{x}_i^T\mathbf{x}_j. \tag{B.27}$$

Taking the difference of the above two equations, we get

$$(\lambda_i - \lambda_j)\,\mathbf{x}_i^T\mathbf{x}_j = 0. \tag{B.28}$$

Since, by hypothesis, $\lambda_i \neq \lambda_j$, it follows that $\mathbf{x}_i^T\mathbf{x}_j = 0$. In vector notation, this is the same as $\mathbf{x}_i\cdot\mathbf{x}_j = 0$. Hence, the eigenvectors $\mathbf{x}_i$ and $\mathbf{x}_j$ are mutually orthogonal.

Suppose that $\lambda_i = \lambda_j = \lambda$. In this case, we cannot conclude that $\mathbf{x}_i^T\mathbf{x}_j = 0$ by the above argument. However, it is easily seen that any linear combination of $\mathbf{x}_i$ and $\mathbf{x}_j$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ. Hence, it is possible to define two new eigenvectors of $\mathbf{A}$, with the eigenvalue λ, which are mutually orthogonal. For instance,

$$\mathbf{x}_i' = \mathbf{x}_i, \tag{B.29}$$

$$\mathbf{x}_j' = \mathbf{x}_j - \left(\frac{\mathbf{x}_i^T\mathbf{x}_j}{\mathbf{x}_i^T\mathbf{x}_i}\right)\mathbf{x}_i. \tag{B.30}$$

It should be clear that this argument can be generalized to deal with any number of eigenvalues which take the same value.

In conclusion, a real symmetric n-dimensional matrix possesses n *real* eigenvalues, with n associated *real* eigenvectors, which are, or can be chosen to be, *mutually orthogonal*.